

Jan Jan 10 08:15:57 2002

GenCore version 4.5  
Copyright (c) 1993 - 2000 Compugen Ltd.

Copyright (c) 1993 - 2000 Compugen Ltd.

OM protein - protein search, using sw model  
January 10, 2002, 02:01:16 : Search time 31.32 Seconds  
(without alignments)  
68,586 Million cell updates/sec

US-09-712-768-8

Title: 142  
1 D5IFLLVCLIRIKGAMSAKOHVFEMA 29

Sequence: BLOSUM62

Scoring table: Gapop 10.0, Gapext 0.5

Searched: 522463 seqs, 74073290 residues 522463

Total number of hits satisfying chosen parameters:

Minimum DB seq length: 0  
Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%  
Maximum Match 100%  
Listing first 45 summaries

Database: A\_Geneseq\_1101.\*  
1: /SID2/gcgdata/geneseq/geneseq/AA1980.DAT.\*  
2: /SID2/gcgdata/geneseq/geneseq/AA1981.DAT.\*  
3: /SID2/gcgdata/geneseq/geneseq/AA1982.DAT.\*  
4: /SID2/gcgdata/geneseq/geneseq/AA1983.DAT.\*  
5: /SID2/gcgdata/geneseq/geneseq/AA1984.DAT.\*  
6: /SID2/gcgdata/geneseq/geneseq/AA1985.DAT.\*  
7: /SID2/gcgdata/geneseq/geneseq/AA1986.DAT.\*  
8: /SID2/gcgdata/geneseq/geneseq/AA1987.DAT.\*  
9: /SID2/gcgdata/geneseq/geneseq/AA1988.DAT.\*  
10: /SID2/gcgdata/geneseq/geneseq/AA1989.DAT.\*  
11: /SID2/gcgdata/geneseq/geneseq/AA1990.DAT.\*  
12: /SID2/gcgdata/geneseq/geneseq/AA1991.DAT.\*  
13: /SID2/gcgdata/geneseq/geneseq/AA1992.DAT.\*  
14: /SID2/gcgdata/geneseq/geneseq/AA1993.DAT.\*  
15: /SID2/gcgdata/geneseq/geneseq/AA1994.DAT.\*  
16: /SID2/gcgdata/geneseq/geneseq/AA1995.DAT.\*  
17: /SID2/gcgdata/geneseq/geneseq/AA1996.DAT.\*  
18: /SID2/gcgdata/geneseq/geneseq/AA1997.DAT.\*  
19: /SID2/gcgdata/geneseq/geneseq/AA1998.DAT.\*  
20: /SID2/gcgdata/geneseq/geneseq/AA1999.DAT.\*  
21: /SID2/gcgdata/geneseq/geneseq/AA2000.DAT.\*  
22: /SID2/gcgdata/geneseq/geneseq/AA2001.DAT.\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

# SUMMARIES

Result No.	Score	Match	Length	DB	ID	Description
1	142	100.0	29	22	AAV97753	G. oxydans cytochr
2	66	46.5	133	21	AA292909	Human secreted pro
3	66	46.5	203	22	AA292909	Gene 26 human secr
4	66	46.5	203	21	AA292909	Human secreted pro
5	66	46.5	203	21	AA292909	Human secreted pro
6	66	46.5	203	21	AA292909	Human secreted pro
7	66	46.5	203	21	AA292909	Human secreted pro
8	66	46.5	203	21	AA292909	Human secreted pro
9	66	46.5	203	21	AA292909	Human secreted pro
10	66	46.5	203	21	AA292909	Human secreted pro
11	66	46.5	203	21	AA292909	Human secreted pro

us-09-712-768-8.rag

12	49	34.5	180	22	AA42392	Human polypeptide
13	46.5	32.7	945	20	AA134694	C. pneumoniae gln
14	46.5	32.4	945	20	AA134694	Porphyromonas gln
15	46.5	32.4	819	20	AA134694	Porphyromonas gln
16	46.5	32.4	819	20	AA134694	Porphyromonas gln
17	46.5	32.4	819	20	AA134694	Porphyromonas gln
18	46.5	32.4	819	20	AA134694	Porphyromonas gln
19	46.5	32.4	819	20	AA134694	Porphyromonas gln
20	46.5	32.4	819	20	AA134694	Porphyromonas gln
21	46.5	32.4	819	20	AA134694	Porphyromonas gln
22	46.5	32.4	819	20	AA134694	Porphyromonas gln
23	46.5	32.4	819	20	AA134694	Porphyromonas gln
24	46.5	32.4	819	20	AA134694	Porphyromonas gln
25	46.5	32.4	819	20	AA134694	Porphyromonas gln
26	46.5	32.4	819	20	AA134694	Porphyromonas gln
27	46.5	32.4	819	20	AA134694	Porphyromonas gln
28	46.5	32.4	819	20	AA134694	Porphyromonas gln
29	46.5	32.4	819	20	AA134694	Porphyromonas gln
30	46.5	32.4	819	20	AA134694	Porphyromonas gln
31	46.5	32.4	819	20	AA134694	Porphyromonas gln
32	46.5	32.4	819	20	AA134694	Porphyromonas gln
33	46.5	32.4	819	20	AA134694	Porphyromonas gln
34	46.5	32.4	819	20	AA134694	Porphyromonas gln
35	46.5	32.4	819	20	AA134694	Porphyromonas gln
36	46.5	32.4	819	20	AA134694	Porphyromonas gln
37	46.5	32.4	819	20	AA134694	Porphyromonas gln
38	46.5	32.4	819	20	AA134694	Porphyromonas gln
39	46.5	32.4	819	20	AA134694	Porphyromonas gln
40	46.5	32.4	819	20	AA134694	Porphyromonas gln
41	46.5	32.4	819	20	AA134694	Porphyromonas gln
42	46.5	32.4	819	20	AA134694	Porphyromonas gln
43	46.5	32.4	819	20	AA134694	Porphyromonas gln
44	46.5	32.4	819	20	AA134694	Porphyromonas gln
45	46.5	32.4	819	20	AA134694	Porphyromonas gln

## ALIGNMENTS

RESULT 1	AAV97753 standard; Protein: 29 AA.
ID	AAV97753
AC	AAV97753
XX	06-ANG-2001 (first entry)
DT	G. oxydans cytochrome C oxidase (COIII) protein sequence fragment.
DE	Cytochrome C oxidase; COI; COII; COIII; COI complex; D-sorbitol;
XX	oxidative fermentation; electron transfer; respiratory chain; L-sorbitol;
KW	2-Keto-L-gluconic acid; 2KGA production; ketone production.
KW	carboxylic acid production; ketone production.
OS	Glucobacter oxydans.
XX	EP1103603-A2.
PN	30-MAY-2001.
XX	14-NOV-2000; 2000EP-0124785.
PD	17-NOV-1999; 99EP-0122842.
PE	(HOF) HOFFMANN LA ROCHE & CO AG F.
XX	Asakura A, Hoshino T, Shinjoh M;
XX	WPI: 2001-357953/38.
XX	N-PSDB: AAA91492.
DR	New cytochrome C oxidase complex having cytochrome C oxidase activity
XX	from Glucobacter C oxidase DSM 4025, useful in mediating electron
PT	

Human polypeptide  
C. pneumoniae gln  
Porphyromonas gln  
Porphyromonas gln  
Amino acid sequenc  
NADL protein encod  
Amino acid sequenc  
Amino acid sequenc  
Neisseria gonorrhoe  
Neisseria meningit  
Sequene encoded b  
Human polypeptide  
Human secreted pro  
Human secreted prote  
Human membrane ass  
Human protein sequ  
Human protein charact  
Protein characteri  
Hog cholesteri virus  
Hog cholesteri virus  
Hog cholesteri virus  
Hog cholesteri virus  
Human secreted prote  
Human secreted prote  
The beta-1,3-galact  
Human beta-1,3-galact  
Arabidopsis thaliana  
C glutamicum prote  
Arabidopsis thaliana  
Arabidopsis thaliana  
Rhesus monkey mela  
Arabidopsis thaliana  
Arabidopsis thaliana

PT transfer in respiratory chain or producing 2-keto-L-gulononic acid from  
 PS L-sorbose or D-sorbitol  
 PS Claim 6; Page 26; 42pp; English.  
 CC This sequence represents a fragment of the gluconohexaric acid from  
 CC cytochrome C oxidase, COII. The invention relates to a cytochrome C  
 CC oxidase COI complex. The COI complex is useful in improving oxidative  
 CC fermentation and is an essential component mediating electron transfer  
 CC in the respiratory chain. The recombinant mediating electron transfer  
 CC cytochrome C oxidase may be used in the genetic preparation of a  
 CC recombinant COI complex and in the production of 2-keto-L-gulononic acid  
 CC acceptor from an enzyme belonging to dehydrogenase C, an electron  
 CC especially the production of 2KGA from L-sorbose or D-sorbitol.  
 SO Sequence 29 AA;

Query Match  
 Best Local Similarity 100.0%; Score 142; DB 22; Length 29;  
 Matches 29; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
 QY 1 DSFLVLCIRIRGMSAKHGFEMAA 29  
 DB 1 dsflvclirirgmsakxhgfemaa 29

RESULT 2  
 ID AAB29909 standard; Protein; 133 AA.  
 AC AAB29909;  
 DT 09-FEB-2001 (first entry)  
 DE Human secreted protein BLAST search protein SEQ ID NO: 167.

CC Cytostatic; immunosuppressive; neutrotropic; neuroprotective; antiviral;  
 CC anti-allergic; hepatotropic; antidiabetic; anti-inflammatory; anticancer;  
 CC vulnarary; anticonvulsant; antibacterial; antitumor; antifungal; anticancer;  
 CC neurologic disease; infection; human; secreted protein.  
 OS Homo sapiens.  
 PN WO200061779-A1.  
 PD 19-OCT-2000.

PF 06-APR-2000; 2000MO-US09068.  
 PR 09-APR-1999; 99US-0128659.  
 PR 20-JAN-2000; 2000US-0177050.  
 PA (HUMA-) HUMAN GENOME SCI INC.  
 PI Rosen CA, Ruben SM, Komatsoulis G.  
 DR WPI; 2000-647424/62.  
 XX Isolated nucleic acid molecule encoding a human secreted protein is  
 XX used in preventing, treating or ameliorating a medical condition -  
 XX Disclosure; Page 487; 495pp; English.

CC The invention relates to the isolation of genes AAC63410-C63458 encoding  
 CC peptide fragments homologous to the protein encoded by the gene given  
 CC in the descriptor line. The sequence is a search result from a BLASTX  
 CC biology search. The genes and proteins are useful for preventing,

CC ameliorating or treating medical conditions, e.g. by protein or gene  
 CC therapy. The genes are isolated from a range of human tissues disclosed  
 CC (ant)agonists are useful in the diagnosis, treatment and prevention of:  
 CC (a) cancer, e.g. breast and ovarian cancer; (b) immune disorders, e.g. Addison's disease,  
 CC allergic, autoimmune hemolytic anaemia, multiple sclerosis, rheumatoid  
 CC arthritis, Crohn's disease, Crohn's disease, Crohn's disease, Crohn's disease,  
 CC as myocardial ischemias; (d) wound healing; (e) neurological diseases  
 CC as viral, bacterial, fungal and parasitic infections.  
 SO Sequence 133 AA;

Query Match  
 Best Local Similarity 46.5%; Score 66; DB 21; Length 133;  
 Matches 14; Conservative 3; Mismatches 11; Indels 0; Gaps 0;  
 QY 2 STFLVLCIRIRGMSAKHGFEMAA 29  
 DB 84 stflvclirirgmsakxhgfemaa 111

RESULT 3  
 ID AAB75603 standard; Protein; 203 AA.  
 AC AAB75603;  
 DT 06-APR-2001 (first entry)  
 DE Gene 46 human secreted protein homologous amino acid sequence #157.

CC Human; secreted protein; immunosuppressive; antiarthritic; antineoplastic;  
 CC antiproliferative; cytosolic; cardiac; vasorelaxant; antidiabetic;  
 CC neurotropic; neuroprotective; antibacterial; vitinoid; fungicide;  
 CC hyperproliferative; vitinoid; antidiabetic; antineoplastic; antineoplastic;  
 CC nervous system disorder; cerebrovascular disorder; wound healing;  
 CC food additive.  
 OS Homo sapiens.  
 PN WO200077026-A1.  
 PD 21-DEC-2000.

PF 01-JUN-2000; 2000MO-US14973.  
 PR 11-JUN-1999; 99US-0138630.  
 PA (HUMA-) HUMAN GENOME SCI INC.  
 PI (ROSE/) ROSEN C A.  
 DR Rosen CA, Ruben SM, Komatsoulis GA.  
 DR WPI; 2001-071258/08.  
 XX Nucleic acid molecules encoding human secreted proteins, used in  
 XX preventing, treating or ameliorating a disorder, e.g. Alzheimer's and  
 XX Parkinson's diseases and cancers -  
 XX Disclosure; Page 94; 542pp; English.

CC Human secreted proteins AAB75506 - AAB75554 are encoded by polynucleotide  
 CC sequences AAF64176 - AAF64224. The specification includes amino acid  
 CC sequences AAB75555 - AAB75606 which represent fragments of the human  
 CC secreted proteins, and protein sequences with which they share human  
 CC activities dependent on the tissues and cells in which they are

xx expressed, examples of these activities include, immunosuppressive;  
 CC antiarthritic; antirheumatic; antiproliferative; cytostatic; cardiant;  
 CC vasotropic; cerebroprotective; nootropic; neuroprotective; antibacterial;  
 CC virocidic; fungicide; ophthalmological; and vulnereary. The proteins,  
 CC polynucleotides, agonists and antagonists can be used to treat or detect  
 CC or diagnose various diseases and disorders including, autoimmune  
 CC diseases e.g. rheumatoid arthritis, hyperproliferative disorders  
 CC e.g. neoplasms of the breast or liver, cardiovascular disorders  
 CC e.g. cardiac arrest, cerebrovascular disorders e.g. cerebral ischemia,  
 CC angiogenesis, nervous system disorders e.g. Alzheimer's disease,  
 CC infections caused by bacteria, viruses and fungi and ocular disorders  
 CC e.g. corneal infection. The polypeptides can also be used to aid wound  
 CC healing and epithelial cell proliferation, to prevent skin ageing due to  
 CC sunburn, to maintain organs before transplantation, for supporting cell  
 CC culture of primary tissues, to regenerate tissues and in chemotaxis. The  
 CC polypeptides can also be used as a food additive or preservative to  
 CC increase or decrease storage capabilities. Included in the invention are  
 CC polynucleotide sequences AAF64167 - AAF64175 and peptide AAB75505 which  
 CC are used in the isolation, identification and characterisation of the  
 CC proteins of the invention.

xx Sequence 203 AA;

xx Query Match 46.5%; Score 66; DB 22; Length 203;  
 xx Best Local Similarity 50.0%; Pred. No. 0.011; Indels 0; Gaps 0;  
 xx Matches 14; Conservative 3; Mismatches 11;

yy 2 SIFLWCLIRIRGAMSAKQHVGFEMAA 29  
 | | | | | : | | | | : | | | | |  
 Db 154 stflavcillrlqghfthsehgfga 181

RESULT 4  
 AAB34657 standard; Protein; 62 AA.

xx AAB34657;

xx 26-JAN-2001 (first entry)

xx Gene 26 human secreted protein homologous amino acid sequence #141.

xx Human; secreted protein; diagnosis; immunosuppressive; antiarthritic;  
 xx antirheumatic; antiproliferative; cytostatic; cardiant; vasotropic;  
 xx cerebroprotective; nootropic; neuroprotective; antibacterial; virocidic;  
 xx fungicide; ophthalmological; gene therapy; autoimmune disease; neoplasm;  
 xx hyperproliferative disorder; cancer; cardiovascular disorder; infection;  
 xx cerebrovascular disorder; angiogenesis; nervous system disorder;  
 xx ocular disorder; wound healing; skin aging; food additive; preservative.

xx Oncorhynchus keta.

xx WO200056751-A1.

xx 28-SEP-2000.

xx 09-MAR-2000; 2000WO-US06013.

xx 19-MAR-1999; 99US-0125360.

xx 11-JUN-1999; 99US-0138626.

xx 03-DEC-1999; 99US-0168662.

xx (HUMA-) HUMAN GENOME SCI INC.

xx Rosen CA, Ruben SM, Komatsoulis G;

xx WPI, 2000-579482/54.

xx Isolated nucleic acid molecule encoding a human secreted protein is  
 xx used in preventing, treating or ameliorating a medical condition  
 xx Disclosure; Page 405; 419pp; English.

xx The polynucleotide sequences given in AAC59738 to AAC59787 encode the  
 CC human secreted proteins given in AAB34577 to AAB34626. AAB34627 to  
 CC AAB34666 represent human secreted polypeptide sequences and proteins  
 CC homologous to them, which are given in the exemplification of the present  
 CC invention. Human secreted proteins have activities based on the tissues  
 CC and cells the genes are expressed in. Example of activities include;  
 CC antiarthritic; immunosuppressive; antirheumatic; antiproliferative;  
 CC cytostatic; cardiant; vasotropic; cerebroprotective; nootropic;  
 CC neuroprotective; antiproliferative; virocidic; fungicide; and  
 CC ophthalmological. The polynucleotides and proteins can be are used to  
 CC prevent, treat or ameliorate a medical condition in e.g. humans, mice,  
 CC rabbits, goats, horses, cats, dogs, chickens or sheep. They are also  
 CC used in diagnosing a pathological condition or susceptibility to a  
 CC pathological condition. Disorders which are diagnosed or treated include  
 CC autoimmune diseases, hyperproliferative disorders e.g. neoplasms and  
 CC cancers of the breast or liver, cardiovascular disorders,  
 CC cerebrovascular disorders, angiogenesis, nervous system disorders,  
 CC infections caused by bacteria, viruses and fungi and ocular disorders.  
 CC The proteins can also be used to aid wound healing and epithelial cell  
 CC proliferation, to prevent skin aging due to sunburn, to maintain organs  
 CC before transplantation, for supporting cell culture of primary tissues,  
 CC to regenerate tissues and in chemotaxis. The proteins can also be used  
 CC as a food additive or preservative to increase or decrease storage  
 CC capabilities. AAC59729 to AAC59737 and AAB34576 represent sequences used  
 CC in the exemplification of the present invention.

xx Sequence 62 AA;

xx Query Match 45.8%; Score 65; DB 21; Length 62;  
 xx Best Local Similarity 46.4%; Pred. No. 0.0042;  
 xx Matches 13; Conservative 6; Mismatches 9; Indels 0; Gaps 0;

yy 2 SIFLWCLIRIRGAMSAKQHVGFEMAA 29  
 | | | | | : | | | | : | | | | |  
 Db 13 stflavcillrlqghfthsehgfga 40

RESULT 5  
 AAB34658 standard; Protein; 62 AA.

xx AAB34658;

xx 26-JAN-2001 (first entry)

xx Human secreted protein sequence encoded by gene 26 SEQ ID NO:142.

xx Human; secreted protein; diagnosis; immunosuppressive; antiarthritic;  
 xx antirheumatic; antiproliferative; cytostatic; cardiant; vasotropic;  
 xx cerebroprotective; nootropic; neuroprotective; antibacterial; virocidic;  
 xx fungicide; ophthalmological; gene therapy; autoimmune disease; neoplasm;  
 xx hyperproliferative disorder; cancer; cardiovascular disorder; infection;  
 xx cerebrovascular disorder; angiogenesis; nervous system disorder;  
 xx ocular disorder; wound healing; skin aging; food additive; preservative.

xx Homo sapiens.

xx WO200056751-A1.

xx 28-SEP-2000.

xx 09-MAR-2000; 2000WO-US06013.

xx 19-MAR-1999; 99US-0125360.

xx 11-JUN-1999; 99US-0138626.

xx 03-DEC-1999; 99US-0168662.

xx (HUMA-) HUMAN GENOME SCI INC.

xx Rosen CA, Ruben SM, Komatsoulis G;

The polynucleotide sequences given in AAC59738 to AAC59787 encode the human secreted proteins given in AAB34577 to AAB34626. AAB34627 to AAB34686 represent human secreted polypeptide sequences and proteins homologous to them, which are given in the exemplification of the present invention. Human secreted proteins have activities based on the tissues and cells the genes are expressed in. Example of activities include: antitumorigenic; immunosuppressive; antirheumatic; antiproliferative; cytostatic; cardiac; vasotropic; cerebroprotective; neurotrophic; cytoprotective; antibacterial; viricide; fungicide; and ophthalmological. The polynucleotides and proteins can be used to prevent, treat or ameliorate a medical condition in e.g. humans, mice, rabbits, goats, horses, cats, dogs, chickens or sheep. They are also used in diagnosing a pathological condition or susceptibility to a pathological condition. Disorders which are diagnosed or treated include autoimmune diseases, hyperproliferative disorders e.g. neoplasms and cancers of the breast or liver, cardiovascular disorders, cerebrovascular disorders, angiogenesis, nervous system disorders, The proteins can also be used to aid wound healing and ocular disorders. Proliferation, to prevent skin aging due to sunburn, to maintain cells before transplantation, for supporting cell culture of primary tissues, as a food additive or preservative in chemotaxis. The proteins can also be used as capillaries. AAC59729 to AAC59737 and AAB34576 represent sequences used in the exemplification of the present invention.

```

QY      2 STELLVCLIRILGAMSAKQHVPEMA 29
      | || : || |
Db      13 stfllcfcfrqllfnftrhfigfaaa 40

```

AC AAB75551;  
XX

Human: Secreted protein

**Homo sapiens.**

MO200077026-A1.

21-DEC-2000.

01-JUN-2000; 2000WO-US14973.

11-JUN-1999; 99US-0138630.

PA (HUMA-) HUMAN GENOME SCI INC  
PA (ROSE/) ROSEN C A.  
XX

XX ruden SM, Komatsoulis GA;

WFI; 2001-071258/08.  
N-PSDB; AAF64221.  
XX

PT Nucleic acid molecules encoding human secreted proteins, used in preventing, treating or ameliorating a disorder, e.g. Alzheimer's and Parkinson's diseases and cancers -  
PT  
XX

Page 490; 542pp; English

Human secreted proteins AAB75506 - AAB75554 are encoded by polynucleotide sequences AAB64176 - AAB64224. The specification includes amino acid sequences AAB75505 - AAB75606 which represent fragments of the human secreted proteins, and protein sequences with which they share homology. The proteins and polynucleotides, their agonists and antagonists have activities dependent on the tissues and cells in which they are expressed, examples of these activities include, immunosuppressive; antithrombotic; antithrombotic; antiproliferative; cytostatic; cardiant; vasotropic; cerebroprotective; neurotropic; neuroprotective; antibacterial; virucide; fungicide; opthalmological; and vulnerary. The proteins, polynucleotides, agonists and antagonists can be used to treat or detect diseases e.g. Rheumatoid arthritis, hyperproliferative disorders e.g. neoplasms of the breast or liver, cardiovascular disorders e.g. cardiac arrest, cerebrovascular disorders e.g. cerebral ischaemia, angioneurotic, nervous system disorders e.g. Alzheimer's disease, infections caused by bacteria, viruses and fungi and ocular disorders e.g. corneal infection. The polypeptides can also be used to aid wound healing and epithelial cell proliferation, to prevent skin ageing due to sunburn, to maintain organs before transplantation, for supporting cell culture of primary tissues, to regenerate tissues and in chemotaxis. The polypeptides can also be used as a food additive or preservative to increase or decrease storage capabilities. Included in the invention are polynucleotide sequences AAB64167 - AAB64175 and peptide AAB75505 which are used in the isolation, identification and characterisation of the proteins of the invention.

Sequence 78 AA;

Query Match	45.1%;	Score 64;	DB 22;	Length 78;
Best Local Similarity	50.0%;	Pred. No. 0.0078;		
Matches 14;	Conservative 3;	Mismatch 1;		

QY	2	STFLVLCYIRILR	SAKQHVGEFMAA	29
			: :	
Db	50	stfltlctfgrqlnftskh	gfyfaaa	77

RESULT  
AAB63840

..... standard; Protein; 93 AA.

20-MAR-2001 (first entry)

Accession	Protein description	Seq ID	NO
Human	breast cancer, gastric	1202	

cytostatic; cancer vaccine.

FD-302 (Rev. 11-27-70)

07 DEC 1963

4b-MAY-2000; 2000WO-US14749





PT Nucleic acid molecules encoding human secreted proteins, used in  
 PT preventing, treating or ameliorating a disorder, e.g. Alzheimer's and  
 PT Parkinson's diseases and cancers -  
 XX  
 PS Disclosure: Page 94; 542pp; English.

CC Human secreted proteins AAB75506 - AAB75554 are encoded by polynucleotide  
 CC sequences AAF64176 - AAF64224. The specification includes amino acid  
 CC sequences AAB75555 - AAB75606 which represent fragments of the human  
 CC secreted proteins, and protein sequences with which they share homology.  
 CC The proteins and polynucleotides, their agonists and antagonists have  
 CC activities dependent on the tissues and cells in which they are  
 CC expressed, examples of these activities include, immunosuppressive;  
 CC antitumor; antineumatic; antiproliferative; cytostatic; cardiant;  
 CC vasotropic; cerebroprotective; neurotropic; neuroprotective; antibacterial;  
 CC virucide; fungicide; opthalmological; and vulnerary. The proteins,  
 CC polynucleotides, agonists and antagonists can be used to treat or detect  
 CC or diagnose various diseases and disorders including, autoimmune  
 CC diseases e.g. rheumatoid arthritis, hyperproliferative disorders  
 CC e.g. neoplasms of the breast or liver, cardiovascular disorders  
 CC e.g. cardiac arrest, cerebrovascular disorders e.g. cerebral ischaemia,  
 CC anglogenesis, nervous system disorders e.g. Alzheimer's disease,  
 CC infections caused by bacteria, viruses and fungi and ocular disorders  
 CC e.g. corneal infection. The polypeptides can also be used to aid wound  
 CC healing and epithelial cell proliferation, to prevent skin ageing due to  
 CC culture of primary tissues, to regenerate tissues and in chemotaxis. The  
 CC polypeptides can also be used as a food additive or preservative to  
 CC increase or decrease storage capabilities. Included in the invention are  
 CC polynucleotide sequences AAF64167 - AAF64175 and peptide AAB75505 which  
 CC are used in the isolation, identification and characterisation of the  
 CC proteins of the invention.

Seq Sequence 203 AA;

Query Match 45.1%; Score 64; DB 22; Length 203;  
 Best Local Similarity 50.0%; Pred. No. 0.022;  
 Matches 14; Conservative 3; Mismatches 11; Indels 0; Gaps 0;

OY 2 SIFLVCLIRLRGMSAKOHGFEMAA 29  
 Db 154 stflctfgrqlfhfctskhfgfetaa 181

RESULT 10

AAB57169  
 ID AAB57169 standard; Protein; 35 AA.

AC AAB57169;

DT 13-MAR-2001 (first entry)

DE Human prostate cancer antigen protein sequence SEQ ID NO:1747.

KW Human; prostate cancer; prostate cancer antigen; detection; diagnosis;  
 KW neuroprotective; cytostatic; cardioactive; immunomodulatory; muscular;  
 KW vulnerary; gastrointestinal; nephrotropic; antiinfective; gynaecological;  
 KW antibacterial; gene therapy; neural; immune; reproductive; renal;  
 KW gastrointestinal; pulmonary; cardiovascular; proliferative disorder;  
 KW wound; infectious disease.

OS Homo sapiens.

PN WO200055174-A1.

PD 21-SEP-2000.

PF 08-MAR-2000; 2000WO-US05988.

PR 12-MAR-1999; 99US-0124270.

PA (HUMA-) HUMAN GENOME SCI INC.

PA (ROSE/) ROSEN C A.  
 XX  
 PI Rosen CA, Ruben SM;  
 XX  
 DR WPI: 2000-587513/55.  
 DR N-PSDB; AAF16372.

PT Prostate cancer associated gene sequences, referred to as prostate  
 PT cancer antigens, useful for treatment, prevention, and diagnosis of  
 PT disorders such as prostate cancer -  
 XX  
 PS Claim 11; Page 2216; 2338pp; English.

CC AAF15566 to AAF15505 encode the human prostate cancer associated  
 CC proteins, called prostate cancer antigens, given in AAB56363 to AAB57302.  
 CC The prostate cancer antigens can have neuroprotective, cytostatic,  
 CC cardioactive, immunomodulatory, muscular, vulnerary, gastrointestinal,  
 CC nephrotropic, antiinfective, gynaecological and antibacterial activities,  
 CC and can be used in gene therapy. The prostate cancer antigen  
 CC polynucleotides may be used for detection of prostate cancer  
 CC identification, as chromosome markers, and for numerous other diagnostic  
 CC or research purposes. The prostate cancer antigens may be used to treat  
 CC disorders such as neural, immune, muscular, reproductive,  
 CC gastrointestinal, pulmonary, cardiovascular, renal, and proliferative  
 CC disorders, wounds, and infectious diseases. AAF15506 to AAF15514 to  
 CC AAB57303 represent sequences used in the exemplification of the present  
 CC invention.

Seq Sequence 35 AA;

Query Match 43.0%; Score 61; DB 21; Length 35;  
 Best Local Similarity 50.0%; Pred. No. 0.0097;  
 Matches 13; Conservative 3; Mismatches 10; Indels 0; Gaps 0;

OY 4 FLVLCIRLRGMSAKOHGFEMAA 29  
 Db 10 fltctfgrqlfhfctskhfgfetaa 35

RESULT 11

AAG75934  
 ID AAG75934 standard; Protein; 93 AA.

AC AAG75934;

DT 03-SEP-2001 (first entry)

DE Human colon cancer antigen protein SEQ ID NO:6698.

KW Human; colon cancer; colon cancer antigen; diagnosis; detection;  
 KW colorectal carcinoma; chromosome 17.

OS Homo sapiens.

PN WO200122920-A2.

PD 05-APR-2001.

PF 28-SEP-2000; 2000WO-US26524.

PR 29-SEP-1999; 99US-0157137.

PR 03-NOV-1999; 99US-0163280.

PA (HUMA-) HUMAN GENOME SCI INC.

PI Ruben SM, Barash SC, Birse CE, Rosen CA;

DR WPI: 2001-235357/24.

DR N-PSDB; AAH35339.

PT Nucleic acids encoding 4277 human colon cancer-associated polypeptides,  
 PT useful for preventing, diagnosing and/or treating colorectal cancers -

XX Claim 11; Page 8164; 9803pp; English.  
 PS  
 XX  
 CC AAH32943 to AAH37195 and AAG73514 to AAG77788 represent human colon  
 CC cancer-associated nucleic acid molecules (N) and proteins (P), where  
 CC the proteins are collectively known as colon cancer antigens. The colon  
 CC cancer antigens have cytostatic activity and can be used in gene  
 CC therapy and vaccine production. N and P may be used in the prevention,  
 CC diagnosis and treatment of diseases associated with inappropriate P  
 CC expression. For example, N and P may be used to treat disorders  
 CC associated with decreased expression by rectifying mutations or deletions  
 CC in a patient's genome that affect the activity of P by expressing  
 CC inactive proteins or to supplement the patients own production of P.  
 CC Additionally, N may be used to produce the colon cancer-associated P,  
 CC by inserting the nucleic acids into a host cell and culturing the cell  
 CC to express the proteins. N and P can be used in the prevention, diagnosis  
 CC and treatment of colorectal carcinomas and cancers. AAH37196 to AAH37204  
 CC and AAB77789 represent sequences used in the exemplification of the  
 CC present invention.  
 CC N.B. Pages 666 to 682 and page 7053 of the sequence listing were  
 CC missing at time of publication, meaning no sequences are present for  
 CC SEQ ID NO:1027 to 1052, 7921 and 7922.  
 CC  
 XX Sequence 93 AA:  
 SQ

Query Match 40.8%; Score 58; DB 22; Length 93;  
 Best Local Similarity 46.4%; Pred. No. 0.085;  
 Matches 13; Conservative 3; Mismatches 12; Indels 0; Gaps 0;

OY 2 SIFLVLCIRILRGAMSAKOHVFEMAA 29  
 DB 66 sftlfcirqlrlfhtskhgfxtaa 93

## RESULT 12

AAH42392 standard; Protein; 180 AA.

AAH42392;

22-OCF-2001 (first entry)

Human polypeptide SEQ ID NO 125.

Human; nootropic; neuroprotective; cytostatic; dermatological; virucide;  
 immunosuppressive; antiinflammatory; anti-HIV; antibacterial; vulnerrary;  
 antiparkinsonian; antischizoid; antianemic; antiallergic; cancer;  
 antineuritic; hepatotropic; cerebroprotective; antiinflammatory;  
 antiallergic; antidiabetic; antilucer; anticonvulsant; antifungal;  
 antiparasitic; cardiac; immune disorder; cardiovascular disorder;  
 neurological disease; infection; nephrotropic; gene therapy; vaccine.

Homo sapiens.

WO200155449-A1.

02-AUG-2001.

17-JAN-2001; 2001MO-US01346.

XX  
 PR 31-JAN-2000; 2000US-0179065.  
 PR 04-FEB-2000; 2000US-0180628.  
 PR 19-MAY-2000; 2000US-0205515.  
 PR 07-JUL-2000; 2000US-0216880.  
 PR 14-JUL-2000; 2000US-0218290.  
 PR 14-AUG-2000; 2000US-0225447.  
 PR 01-SEP-2000; 2000US-0229343.  
 PR 06-SEP-2000; 2000US-0230437.  
 PR 08-SEP-2000; 2000US-0231243.  
 PR 25-SEP-2000; 2000US-0234997.  
 PR 29-SEP-2000; 2000US-0236367.  
 PR 13-OCT-2000; 2000US-0239937.  
 XX

PR 08-NOV-2000; 2000US-0246476.  
 PR 08-NOV-2000; 2000US-0246477.  
 PR 08-NOV-2000; 2000US-0246525.  
 PR 08-NOV-2000; 2000US-0246526.  
 PR 08-NOV-2000; 2000US-0246527.  
 PR 17-NOV-2000; 2000US-0249210.  
 PR 17-NOV-2000; 2000US-0249211.  
 PR 17-NOV-2000; 2000US-0249214.  
 PR 17-NOV-2000; 2000US-0249215.  
 PR 01-DEC-2000; 2000US-0250160.  
 PR 01-DEC-2000; 2000US-0250391.  
 PR 05-DEC-2000; 2000US-0251030.  
 PR 05-DEC-2000; 2000US-0251988.  
 PR 05-DEC-2000; 2000US-0256719.  
 PR 06-DEC-2000; 2000US-0251479.  
 PR 08-DEC-2000; 2000US-0251989.  
 PR 08-DEC-2000; 2000US-0251990.  
 PR 11-DEC-2000; 2000US-0254097.  
 XX

(HUMA-) HUMAN GENOME SCI INC.

Rosen CA, Barash SC, Ruben SM;

WPI; 2001-476225/51.

N-PDB; AAI62797.

Novel plasma membrane associated proteins useful for diagnosing,  
 treating, preventing and/or prognosing disorders related to the  
 proteins, including cancer, immune response and neuronal disorders

Claim 11; SEQ ID NO 125; 532pp + Sequence Listing; English.

The invention relates to novel genes (AAI62752-AAI62961) and proteins  
 (AAH42347-AAH42415) useful for preventing, treating or ameliorating  
 medical conditions e.g. by protein or gene therapy. The genes are  
 isolated from a range of human tissues disclosed in the specification.  
 The nucleic acids, proteins, antibodies and (ant)agonists are useful  
 in the diagnosis, treatment and prevention of: (a) cancer, e.g. breast  
 and ovarian cancer and other cancers of the adrenal gland, bone, bone  
 marrow, breast, gastrointestinal tract, liver, lung, or urogenital;  
 (b) immune disorders e.g. Addison's disease, allergies, autoimmune  
 haemolytic anaemia, autoimmune thyroiditis, diabetes mellitus, Crohn's  
 disease, multiple sclerosis, rheumatoid arthritis and ulcerative  
 colitis; (c) cardiovascular disorders such as myocardial ischaemias;  
 (d) wound healing; (e) neurological diseases e.g. cerebral anoxia and  
 epilepsy; and (f) infectious diseases such as viral, bacterial, fungal  
 and parasitic infections.  
 Note: The sequence data for this patent did not form part of the  
 printed specification, but was obtained in electronic format directly  
 from WIPO at ftp.wipo.int/pub/published\_pct\_sequences.

Sequence 180 AA:

Query Match 34.5%; Score 49; DB 22; Length 180;  
 Best Local Similarity 27.3%; Pred. No. 4.7;  
 Matches 6; Conservative 9; Mismatches 7; Indels 0; Gaps 0;

OY 2 SIFLVLCIRILRGAMSAKOHV 23  
 DB 127 avflvvcvnmfkgrqnehl 148

## RESULT 13

AAH34694 standard; Protein; 945 AA.

AAH34694;

13-SEP-1999 (first entry)

C. pneumoniae protein involved in metabolism of polypeptides.

KW Respiratory disease; pneumonia; bronchitis; heart disease; sarcoidosis;  
 KW sinusitis; purulent otitis media; erythema nodosum; pharyngitis;  
 KW vaccine; neutralising epitope.  
 XX  
 XX Chlamydia pneumoniae.  
 OS  
 PN WO927105-A2.  
 XX  
 PD 03-JUN-1999.  
 XX  
 PF 20-NOV-1998; 98WO-1B01890.  
 XX  
 PR 04-NOV-1998; 98US-0107078.  
 PR 21-NOV-1997; 97FR-0014673.  
 XX  
 PA (GEST ) GENSET.  
 XX  
 PI Griffais R;  
 XX  
 DR WPI: 1999-357842/30.  
 XX  
 PT Genome sequence of Chlamydia pneumoniae  
 XX  
 PS Page 692-694; Disclosure: 1912pp; English.  
 XX  
 CC AAY34584-Y35879 represent the proteins encoded by all the open reading  
 CC frames in the complete genome (see AAX91990) of Chlamydia pneumoniae.  
 CC C. pneumoniae causes respiratory disease such as pneumonia and  
 CC bronchitis and is thought to be a contributing factor in heart  
 CC disease, sarcoidosis, sinusitis, purulent otitis media, erythema  
 CC nodosum or pharyngitis. The polypeptides encoded by the open reading  
 CC frames of the C. pneumoniae genome (see AAY34584-Y35879) can be used in  
 CC immunogenic compositions as vaccines. Vectors containing C. pneumoniae  
 CC nucleotide sequences can also be used as immunogenic compositions,  
 CC especially where the vector directs the expression of a neutralising  
 CC epitope of C. pneumoniae.  
 CC  
 XX  
 SQ Sequence 945 AA;  
 QY  
 Db 855 lfvlgyctalllgaaekqh 875  
 QY 3 IFLLV-CLIRIRGAMSAKH 22  
 :|:| | :|:|:| |  
 Query Match 32.7%; Score 46.5; DB 20; Length 945;  
 Best Local Similarity 42.9%; Pred. No. 73;  
 Matches 9; Conservative 7; Mismatches 4; Indels 1; Gaps 1;

RESULT 14  
 AAY34530  
 ID AAY34530 standard; Protein; 750 AA.  
 XX  
 AC AAY34530;  
 XX  
 DT 25-AUG-1999 (first entry)  
 XX  
 DE Porphyromonas gingivalis protein PG61.  
 XX  
 KW Porphyromonas gingivalis; PG; periodontal disease; gingivitis;  
 KW vaccine; antigenic.  
 XX  
 OS Porphyromonas gingivalis.  
 XX  
 PN WO929870-A1.  
 XX  
 PD 17-JUN-1999.  
 XX  
 PF 10-DEC-1998; 98WO-AU01023.  
 XX  
 PR 04-AUG-1998; 98AU-0005028.  
 PR 10-DEC-1997; 97AU-0000839.  
 PR 31-DEC-1997; 97AU-0001182.

PR 30-JAN-1998; 98AU-0001546.  
 PR 10-MAR-1998; 98AU-0002264.  
 PR 09-APR-1998; 98AU-0002911.  
 PR 23-APR-1998; 98AU-0003128.  
 PR 03-MAY-1998; 98AU-0003338.  
 PR 22-MAY-1998; 98AU-0003654.  
 PR 29-JUL-1998; 98AU-0004917.  
 XX  
 PA (CSLC-) CSL LTD.  
 XX  
 PI Agius CT, Barr IG, Hocking DM, Margetts MB, Patterson MA;  
 PI Ross BC, Rothel LJ, Webb EA;  
 XX  
 DR WPI: 1999-385613/32.  
 DR N-PSDB: AAX91748.  
 XX  
 PT Antigenic Porphyromonas gingivalis peptides for preventing  
 PT gingivitis  
 XX  
 PS Claim 1; Page 520-521; 588pp; English.  
 XX  
 CC AAX91536 to AAX91801 encode two hundred and sixty six antigenic  
 CC Porphyromonas gingivalis (PG) polypeptide sequences given in AAY34318 to  
 CC AAY34583. AAX91802 to AAX91989 represent PCR primers used in the  
 CC isolation of the PG polypeptides. The PG polypeptides have antibacterial  
 CC activity with a vaccine mechanism of action. The PG polypeptides can be  
 CC used as vaccines especially against Porphyromonas gingivalis. Probes can  
 CC be used to detect Porphyromonas gingivalis in standard hybridisation  
 CC assays. Porphyromonas gingivalis is involved in periodontal disease  
 CC especially gingivitis.  
 CC  
 XX  
 SQ Sequence 750 AA;  
 QY  
 Db 635 vlagemlclshagfema 651  
 QY 12 ILRGAMSAKHVGFEWA 28  
 :|:| | |  
 Query Match 32.4%; Score 46; DB 20; Length 750;  
 Best Local Similarity 47.1%; Pred. No. 68;  
 Matches 8; Conservative 3; Mismatches 6; Indels 0; Gaps 0;

RESULT 15  
 AAY34397  
 ID AAY34397 standard; Protein; 819 AA.  
 XX  
 AC AAY34397;  
 XX  
 DT 25-AUG-1999 (first entry)  
 XX  
 DE Porphyromonas gingivalis protein PG61.  
 XX  
 KW Porphyromonas gingivalis; PG; periodontal disease; gingivitis;  
 KW vaccine; antigenic.  
 XX  
 OS Porphyromonas gingivalis.  
 XX  
 PN WO929870-A1.  
 XX  
 PD 17-JUN-1999.  
 XX  
 PF 10-DEC-1998; 98WO-AU01023.  
 XX  
 PR 04-AUG-1998; 98AU-0005028.  
 PR 10-DEC-1997; 97AU-0000839.  
 PR 31-DEC-1997; 97AU-0001182.  
 PR 30-JAN-1998; 98AU-0001546.  
 PR 10-MAR-1998; 98AU-0002264.  
 PR 09-APR-1998; 98AU-0002911.  
 PR 23-APR-1998; 98AU-0003128.  
 PR 25-MAY-1998; 98AU-0003338.  
 PR 22-MAY-1998; 98AU-0003654.

```

PR 29-JUL-1998; 98AU-0004917.
XX
PA (CSLC-) CSL LTD.
XX
PI Agius CT, Barr IG, Hocking DM, Margetts MB, Patterson MA;
PI Ross BC, Rochel LJ, Webb EA;
XX
DR WPI; 1999-385613/32.
DR N-PSDB; AAX91615.
XX
PT Antigenic Porphyromonas gingivalis peptides for preventing
PT gingivitis
XX
PS Claim 1; Page 367-369; 588pp; English.
XX
CC AAX91536 to AAX91801 encode two hundred and sixty six antigenic
CC Porphyromonas gingivalis (PG) polypeptide sequences given in AAY34318 to
CC AAY34583. AAX91802 to AAX91989 represent PCR primers used in the
CC isolation of the PG polypeptides. The PG polypeptides have antibacterial
CC activity with a vaccine mechanism of action. The PG polypeptides can be
CC used as vaccines especially against Porphyromonas gingivalis. Probes can
CC be used to detect Porphyromonas gingivalis in standard hybridisation
CC assays. Porphyromonas gingivalis is involved in periodontal disease
CC especially gingivitis.
XX
SQ Sequence 819 AA;

```

Query Match 32.4%; Score 46; DB 20; Length 819;  
 Best Local Similarity 47.1%; Pred. No. 75;  
 Matches 8; Conservative 3; Mismatches 6; Indels 0; Gaps 0;

```

OY 12 ILRGAMSAKOHVGFEMA 28
   :|:| | ||||
Db 704 vlagsmllshagfema 720

```

Search completed: January 10, 2002, 02:01:17  
 Job time: 290 sec







Db 212 STELTICINOLMEHTSKHHFEEAA 239

RESULT 2

; Sequence 2, Application US/08750717

; Patent No. 6180109

; GENERAL INFORMATION:

; APPLICANT: MOORMANN, Robertus J. M.

; APPLICANT: VAN RIJN, Petrus A.

; TITLE OF INVENTION: Nucleotide Sequences of Pestivirus

; TITLE OF INVENTION: Strains, Polypeptides Encoded by These Sequences and Use

; TITLE OF INVENTION: Theoret for Diagnosis and Prevention of Pestivirus

; NUMBER OF SEQUENCES: 11

; CORRESPONDENCE ADDRESS:

; ADDRESSEE: YOUNG & THOMPSON

; STREET: 745 South 23rd Street

; CITY: Arlington

; STATE: Virginia

; COUNTRY: USA

; ZIP: 22202

; COMPUTER READABLE FORM:

; MEDIUM TYPE: Floppy disk

; OPERATING SYSTEM: IBM PC compatible

; SOFTWARE: Patentin Release #1.0, Version #1.30

; CURRENT APPLICATION DATA:

; APPLICATION NUMBER: US/08/750,717

; FILING DATE: 24-DEC-1996

; CLASSIFICATION: 424

; PRIOR APPLICATION DATA:

; APPLICATION NUMBER: EP 94201743.5

; FILING DATE: 17-JUN-1994

; PRIOR APPLICATION DATA:

; APPLICATION NUMBER: WO PCT/NL95/00214

; FILING DATE: 16-JUN-1995

; ATTORNEY/AGENT INFORMATION:

; NAME: PATCH, Andrew J.

; REGISTRATION NUMBER: 32,925

; REFERENCE/DOCKET NUMBER: BO 39123

; TELECOMMUNICATION INFORMATION:

; TELEPHONE: 703-521-2297

; TELEFAX: 703-685-0573

; TELEX: 248425 EMBON

; INFORMATION FOR SEQ ID NO: 2:

; SEQUENCE CHARACTERISTICS:

; LENGTH: 3898 amino acids

; TYPE: amino acid

; TOPOLOGY: linear

; MOLECULE TYPE: protein

; US-08-750-717-2

Query Match 33.1%; Score 47; DB 4; Length 3898;

Best Local Similarity 57.1%; Pred. No. 91;

Matches 8; Conservative 4; Mismatches 2; Indels 0; Gaps 0;

QY 2 STELVCLIRLNG 15

Db 659 TIAFLICLIRKVRG 672

RESULT 3

; US-08-876-991-2

; Sequence 2, Application US/08876991

; Patent No. 5925360

; GENERAL INFORMATION:

; APPLICANT: Gregor Meyers, Tillmann R menapf,

; APPLICANT: Heinz-J rgen Thiel

; TITLE OF INVENTION: Hog cholera virus vaccine and diagnostic

; NUMBER OF SEQUENCES: 13

; CORRESPONDENCE ADDRESS:

; ADDRESSEE: Organon Teknika Corporation

; ADDRESSEE: Biotechnology Research Institute

; STREET: 1330-A Piccard Drive

; CITY: Rockville

; STATE: Maryland

; COUNTRY: U.S.A.

; ZIP: 20850

; COMPUTER READABLE FORM:

; MEDIUM TYPE: Floppy disk

; OPERATING SYSTEM: IBM PC compatible

; SOFTWARE: Patentin Release #1.0, Version #1.25

; CURRENT APPLICATION DATA:

; APPLICATION NUMBER: US/08/876,991

; FILING DATE: 16-JUN-1997

; CLASSIFICATION: 424

; PRIOR APPLICATION DATA:

; APPLICATION NUMBER: US/08/747,577

; FILING DATE:

; APPLICATION NUMBER: US/08/650,584

; FILING DATE:

; APPLICATION NUMBER: US/08/469,702

; FILING DATE:

; APPLICATION NUMBER: US/08/123,596

; FILING DATE:

; APPLICATION NUMBER: 07/797,554

; FILING DATE: 22-NOV-1991

; APPLICATION NUMBER: US 07/494,991

; FILING DATE: 16-MAR-1990

; CLASSIFICATION: 424

; ATTORNEY/AGENT INFORMATION:

; NAME: William M. Blackstone

; REGISTRATION NUMBER: 29,772

; REFERENCE/DOCKET NUMBER:

; TELECOMMUNICATION INFORMATION:

; TELEPHONE: (301) 258-5200

; INFORMATION FOR SEQ ID NO: 2:

; SEQUENCE CHARACTERISTICS:

; LENGTH: 3898 amino acids

; TYPE: amino acid

; TOPOLOGY: linear

; MOLECULE TYPE: protein

; US-08-876-991-2

Query Match 31.0%; Score 44; DB 2; Length 3898;

Best Local Similarity 70.0%; Pred. No. 2.8e+02;

Matches 7; Conservative 3; Mismatches 0; Indels 0; Gaps 0;

QY 6 LVCLIRLNG 15

Db 663 LACLIKVRG 672

RESULT 4

; US-09-059-853-2

; Sequence 2, Application US/09059853

; Patent No. 5935582

; GENERAL INFORMATION:

; APPLICANT: Gregor Meyers, Tillmann R menapf,

; APPLICANT: Heinz-J rgen Thiel

; TITLE OF INVENTION: Hog cholera virus vaccine and diagnostic

; NUMBER OF SEQUENCES: 13

; CORRESPONDENCE ADDRESS:

; ADDRESSEE: Organon Teknika Corporation

; ADDRESSEE: Biotechnology Research Institute

; STREET: 1330-A Piccard Drive

; CITY: Rockville

; STATE: Maryland

; COUNTRY: U.S.A.

; ZIP: 20850

; COMPUTER READABLE FORM:

; MEDIUM TYPE: Floppy disk

```

;
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patentin Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/09/059,853
; FILING DATE:
; CLASSIFICATION:
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 07/97,554
; FILING DATE: 22-NOV-1991
; APPLICATION NUMBER: US 07/494,991
; FILING DATE: 16-MAR-1990
; ATTORNEY/AGENT INFORMATION:
; NAME: William M. Blackstone
; REGISTRATION NUMBER: 29,772
; REFERENCE/DOCKET NUMBER:
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (301) 258-5200
; INFORMATION FOR SEQ ID NO: 2:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 3898 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: protein
; US-09-059-853-2

```

```

Query Match      31.0%; Score 44; DB 2; Length 3898;
Best Local Similarity 70.0%; Pred. No. 2.8e+02;
Matches 7; Conservative 3; Mismatches 0; Indels 0; Gaps 0;

```

OY 6 LVCLIRILRG 15  
 Db 663 LVCLIRKVRG 672

```

RESULT 5
US-08-976-255-14
; Sequence 14, Application US/08976255
; Patent No. 6136581
; GENERAL INFORMATION:
; APPLICANT: Jono, Keith E.
; TITLE OF INVENTION: KINASE GENES AND USES
; NUMBER OF SEQUENCES: 53
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Lyon & Lyon
; STREET: 633 West Fifth Street
; STREET: Suite 4700
; CITY: Los Angeles
; STATE: California
; COUNTRY: U.S.A.
; ZIP: 90071-2066
; COMPUTER READABLE FORM:
; MEDIUM TYPE: 3.5" Diskette, 1.44 MB
; OPERATING SYSTEM: IBM P.C. DOS 5.0
; SOFTWARE: FASTSEQ for Windows 2.0
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/976,255
; FILING DATE: No. 6136581ember 21, 1997
; CLASSIFICATION: 435
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 60/031,675
; FILING DATE: No. 6136581ember 22, 1996
; ATTORNEY/AGENT INFORMATION:
; NAME: Warburg, Richard J.
; REGISTRATION NUMBER: 32,327
; REFERENCE/DOCKET NUMBER: 229/182
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (213) 489-1600
; TELEFAX: (213) 955-0440

```

```

;
; TELEX: 67-3510
; INFORMATION FOR SEQ ID NO: 14:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 1503 amino acids
; TYPE: amino acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: Protein
; US-08-976-255-14

```

```

Query Match      29.6%; Score 42; DB 4; Length 1503;
Best Local Similarity 35.0%; Pred. No. 2e+02;
Matches 7; Conservative 8; Mismatches 5; Indels 0; Gaps 0;

```

OY 4 FLVCLIRILRGAMSAKQHV 23  
 Db 723 FLVCLIRILRGAMSAKQHV 742

```

RESULT 6
5206163-1
; Patent No. 5206163
; APPLICANT: RENARD, ANDRE;DINA, DINO;MARTIAL, JOSEPH
; TITLE OF INVENTION: DNA ENCODING BOVINE DIARRHEA
; VIRUS PROTEIN
; NUMBER OF SEQUENCES: 3
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/07/550,816
; FILING DATE: 06-JUL-1990
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 331,037
; FILING DATE: 29-MAR-1989
; APPLICATION NUMBER: 752,981
; FILING DATE: 08-JUL-1985
; SEQ ID NO: 1:
; LENGTH: 1286
; 5206163-1

```

```

Query Match      28.9%; Score 41; DB 6; Length 1286;
Best Local Similarity 60.0%; Pred. No. 2.5e+02;
Matches 6; Conservative 4; Mismatches 0; Indels 0; Gaps 0;

```

OY 6 LVCLIRILRG 15  
 Db 666 LVCLIRKVRG 675

```

RESULT 7
US-08-786-748A-38
; Sequence 38, Application US/08786748A
; Patent No. 5714577
; GENERAL INFORMATION:
; APPLICANT: Ronald, Montelaro C.
; APPLICANT: Tencza, Sarah B.
; APPLICANT: Mletzner, Timothy A.
; TITLE OF INVENTION: NOVEL ANTIMICROBIAL PEPTIDES
; NUMBER OF SEQUENCES: 169
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Brumbaugh, Graves, Donohue & Raymond
; STREET: 30 Rockefeller Plaza
; CITY: New York
; STATE: New York
; COUNTRY: USA
; ZIP: 10112-0228
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Diskette
; OPERATING SYSTEM: IBM Compatible
; SOFTWARE: FASTSEQ Version 2.0
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/786,748A

```

TYPE: amino acid  
STRANDEDNESS: single  
TOPOLOGY: linear

MOLECULE TYPE  
US-08-932-682-38

```

22Y      9 LIRILRGAMSAKHV 23
          :|:|:| | | :|:
Db       2 VIRVVRGACRAIRHI 16

```

## RESULT 10

US-08-932-682-39  
Sequence 39, Application US/08932682  
Patent No. 5945507  
GENERAL INFORMATION:



LENGTH: 236 amino acids  
TYPE: amino acid  
STRANDEDNESS: single  
TOPOLOGY: linear  
US-08-460-746A-3

Query Match 28.2%; Score 40; DB 1; Length 236;  
Best Local Similarity 40.0%; Pred. No. 52;  
Matches 8; Conservative 5; Mismatches 7; Indels 0; Gaps 0;

QY 3 IFLVCLIRLRGMSAKOH 22  
:||||: :| :||:|  
Db 10 VLLVCVAVALSASAEPH 29

RESULT 13  
US-08-460-555-3  
Sequence 3, Application US/08460555  
Patent No. 5827671  
GENERAL INFORMATION:  
APPLICANT: MATLASHESKI, Gregory  
TITLE OF INVENTION: DIFFERENTIALLY EXPRESSED LEISHMANIA  
NUMBER OF SEQUENCES: 4  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: Sim & McBurney  
STREET: Suite 701, 330 University Avenue  
CITY: Toronto  
STATE: Ontario  
COUNTRY: Canada  
ZIP: M5G 1R7  
COMPUTER READABLE FORM:  
MEDIUM TYPE: Floppy disk  
COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: Patentln Release #1.0, Version #1.25  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/08/460,555  
FILING DATE:  
CLASSIFICATION: 435  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: US 08/302,463  
FILING DATE: 12-SEP-1994  
CLASSIFICATION: 435  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: US 08/115,987  
FILING DATE: 03-SEP-1993  
CLASSIFICATION: 435  
ATTORNEY/AGENT INFORMATION:  
NAME: STEWART, Michael I.  
REGISTRATION NUMBER: 24,973  
REFERENCE/DOCKET NUMBER: 1038-485 MIS:V9  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: (416) 595-1155  
TELEFAX: (416) 595-1163  
INFORMATION FOR SEQ ID NO: 3:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 236 amino acids  
TYPE: amino acid  
STRANDEDNESS: single  
TOPOLOGY: linear  
US-08-460-555-3

Query Match 28.2%; Score 40; DB 2; Length 236;  
Best Local Similarity 40.0%; Pred. No. 52;  
Matches 8; Conservative 5; Mismatches 7; Indels 0; Gaps 0;  
QY 3 IFLVCLIRLRGMSAKOH 22  
:||||: :| :||:|

Db 10 VLLVCVAVALSASAEPH 29

RESULT 14  
US-08-460-066-3  
Sequence 3, Application US/08460066  
Patent No. 6133017  
GENERAL INFORMATION:  
APPLICANT: MATLASHESKI, Gregory  
TITLE OF INVENTION: DIFFERENTIALLY EXPRESSED LEISHMANIA  
NUMBER OF SEQUENCES: 4  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: Sim & McBurney  
STREET: Suite 701, 330 University Avenue  
CITY: Toronto  
STATE: Ontario  
COUNTRY: Canada  
ZIP: M5G 1R7  
COMPUTER READABLE FORM:  
MEDIUM TYPE: Floppy disk  
COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: Patentln Release #1.0, Version #1.25  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/08/460,066  
FILING DATE:  
CLASSIFICATION: 435  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: US 08/302,463  
FILING DATE: 12-SEP-1994  
CLASSIFICATION: 435  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: US 08/115,987  
FILING DATE: 03-SEP-1993  
CLASSIFICATION: 435  
ATTORNEY/AGENT INFORMATION:  
NAME: STEWART, Michael I.  
REGISTRATION NUMBER: 24,973  
REFERENCE/DOCKET NUMBER: 1038-484 MIS:V9  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: (416) 595-1155  
TELEFAX: (416) 595-1163  
INFORMATION FOR SEQ ID NO: 3:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 236 amino acids  
TYPE: amino acid  
STRANDEDNESS: single  
TOPOLOGY: linear  
US-08-460-066-3

Query Match 28.2%; Score 40; DB 4; Length 236;  
Best Local Similarity 40.0%; Pred. No. 52;  
Matches 8; Conservative 5; Mismatches 7; Indels 0; Gaps 0;

QY 3 IFLVCLIRLRGMSAKOH 22  
:||||: :| :||:|  
Db 10 VLLVCVAVALSASAEPH 29

RESULT 15  
US-08-035-928-2  
Sequence 2, Application US/08035928  
Patent No. 5538844  
GENERAL INFORMATION:  
APPLICANT: Duyao, Mabel P.  
APPLICANT: MacDonald, Marcy E.  
APPLICANT: Gusella, James F.  
TITLE OF INVENTION: A No. 5538844el Transport Protein Gene from  
the Huntington's Disease Region  
NUMBER OF SEQUENCES: 21



CORRESPONDENCE ADDRESS:

ADDRESSEE: Sterne, Kessler, Goldstein & Fox  
STREET: 1225 Connecticut Avenue N.W.  
CITY: Washington

STATE: D.C.  
COUNTRY: U.S.A.  
ZIP: 20036

COMPUTER READABLE FORM:

MEDIUM TYPE: Floppy disk  
COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: Patent Release #1.0, Version #1.25  
CURRENT APPLICATION DATA:

APPLICATION NUMBER: US/08/035,928  
FILING DATE: 19930323

CLASSIFICATION: 435  
TELECOMMUNICATION INFORMATION:

TELEPHONE: (202) 466-0800  
TELEFAX: (202) 833-8716

TELEX:

INFORMATION FOR SEQ ID NO: 2:

SEQUENCE CHARACTERISTICS:

LENGTH: 455 amino acids  
TYPE: AMINO ACID  
TOPOLOGY: linear  
MOLECULE TYPE: protein

US-08-035-928-2

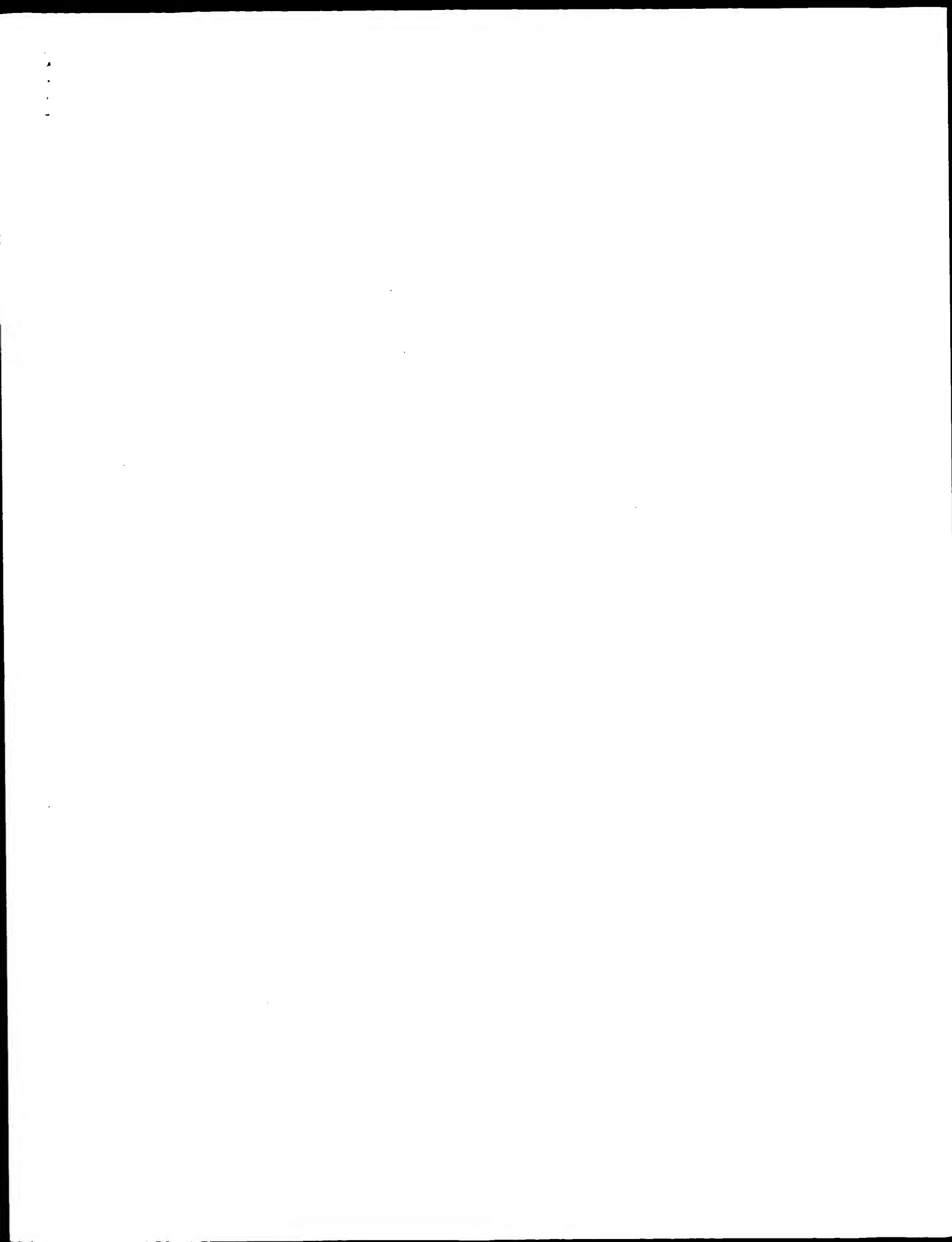
Query Match

Best Local Similarity 28.2%; Score 40; DB 1; Length 455;  
Matches 9; Conservative 7; Mismatches 10; Indels 0; Gaps 0;

Qy 4 FLVCLIRIRGMSAKOHVGFEMAA 29

Db 313 FLIGLTMATIGCAVARIRHPGGEVNA 338

Search completed: January 10, 2002, 02:01:57  
Job time: 279 sec



/cg2_6/pdata/2/ina/5A.COMB.seq:US-08-470-950-1-	44.50	105.44	318.61	2360
/cg2_6/pdata/2/ina/5A.COMB.seq:US-08-467-781-1-	44.50	105.44	318.61	2360
/cg2_6/pdata/2/ina/5A.COMB.seq:US-08-195-987-1-	44.50	105.44	318.61	2360
/cg2_6/pdata/2/ina/5B.COMB.seq:US-08-483-924-1-	44.50	105.44	318.61	2360
/cg2_6/pdata/2/ina/PCYUS.COMB.seq:PCY-US93-06160-1-	44.50	105.44	318.61	2360

```

seq-name: /cygdrive/p/odata/4/ind/0B_comp.seq:US-09-020-956-6
seq-documentation_block:
: Sequence 6, Application US/09020956
: Patent No. 6261562
:
: GENERAL INFORMATION:
:
: APPLICANT: Xu, JIANGCHUN
: APPLICANT: DILLIN, DAVIN C.
: TITLE OF INVENTION: COMPOUNDS FOR IMMUNOTHERAPY OF PROSTATE CANCER AND METHODS
: NUMBER OF SEQUENCES: 178
: CORRESPONDENCE ADDRESS:
: ADDRESSEE: SEED AND BERRY LLP
: STREET: 6300 Columbia Center, 701 Fifth Avenue
: CITY: Seattle

```

COUNTRY: USA  
ZIP: 98104  
COMPUTER READABLE FORM:  
MEDIUM type: Floppy disk  
COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: PatentIn Release #1.0, Version #1.30

CLASSIFICATION:

```

ATTORNEY/AGENT INFORMATION:
NAME: MAKI, David J.
REGISTRATION NUMBER: 31,392
REFERENCE/DOCKET NUMBER: 210121.427C2
TELECOMMUNICATION INFORMATION:
TELEPHONE: (206) 622-4900
TELEFAX: (206) 682-6031
INFORMATION FOR SEQ ID NO: 6:
SEQUENCE CHARACTERISTICS:
LENGTH: 818 base pairs
TYPE: nucleic acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: cDNA
US-09-020-956-6

alignment_scores:
Quality: 64.00 Length: 28
Ratio: 3.200 Gaps: 0
Percent Similarity: 71.429 Percent Identity: 50.000

alignment_block:
US-09-712-768-8 x US-09-020-956-6/rev ..

Align seg 1/1 to reverse of: US-09-020-956-6 from: 1 to: 818

2 SerIlePheLeuLeuValCysLeuIleArgIleLeuArgGlyAlaMetse 18
||| ||||| ::|||::||| ||| ::|||
171 TCAACTTTCCTCACATCTGCTTCATCCGCACACATATATTCACTTAC 122
18 fAlAlysGlnHisValGlyPheGluMetAlaAla 29
:::||||::||| ||||| |||||
121 ATCCAAACATCATCTTGCTTCGGAAGCCGCCGC 88

seq_name: /cgn2_6/prodata/2/ina/5b_COMB.seq:US-09-030-607-6

seq documentation block:
; Sequence 6, Application US/09030607
; Patent No. 6262245
; GENERAL INFORMATION:

```

APPLICANT: Xu, Jiangchun  
APPLICANT: Dillon, David C.  
TITLE OF INVENTION: COMPOUNDS FOR IMMUNOTHERAPY OF PROSTATE CANCER AND METHODS FO  
NUMBER OF SEQUENCES: 224  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: SEED and BERRY LLP  
STREET: 6300 Columbia Center, 701 Fifth Avenue  
CITY: Seattle  
STATE: WA  
COUNTRY: USA  
ZIP: 98104  
COMPUTER READABLE FORM:  
MEDIUM TYPE: Floppy disk  
COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: Patentin Release #1.0, Version #1.30  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/09/030,607  
FILING DATE: 25-FEB-1998  
CLASSIFICATION:  
ATTORNEY/AGENT INFORMATION:  
NAME: MAKI, David J.  
REGISTRATION NUMBER: 31,392  
REFERENCE/DOCKET NUMBER: 210121.427C3  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: (206) 682-6031  
INFORMATION FOR SEQ ID NO: 6:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 818 base pairs  
TYPE: nucleic acid  
STRANDEDNESS: single  
TOPOLOGY: linear  
MOLECULE TYPE: CDNA  
US-09-030-607-6

alignment\_scores:  
Quality: 64.00 Length: 28  
Ratio: 3.200 Gaps: 0  
Percent Similarity: 71.429 Percent Identity: 50.000

alignment\_block:  
US-09-712-768-8 x US-09-030-607-6/rev ..

Align seg 1/1 to reverse of: US-09-030-607-6 from: 1 to: 818

```
2 SerTlePheLeuLeuValCysLeuIleArgIleLeuArgGlyAlaMetse 18  
||| ||||| ::|||::||| ||| ::|||  
171 TCACTTTCCTCCTCATCTGCTTCATCCGCCAACATAATATTTCACCTTAC 122  
18 rAlaLysGlnHisValGlyPheGluMetAlaIa 29  
:::|||||::||| ||||| |||||  
121 ATCCAAACATCATCTTGGCTTGAAGCGCCGCC 88
```

seq\_name: /cgn2\_6/ptodata/2/1na/5A\_COMB.seq:US-08-219-842-3

seq\_documentation\_block:

Sequence 3, Application US/08219842  
Patent No. 5565323  
GENERAL INFORMATION:  
APPLICANT: Parker, W. D.  
APPLICANT: Herrnstadt, Corinna  
TITLE OF INVENTION: Diagnostic and Therapeutic Compositions  
NUMBER OF SEQUENCES: 95  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: Campbell and Flores  
STREET: 4370 La Jolla Village Drive, Suite 700  
CITY: San Diego  
STATE: California  
COUNTRY: USA  
ZIP: 92122

COMPUTER READABLE FORM:  
MEDIUM TYPE: Floppy disk  
COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: Patentin Release #1.0, Version #1.25  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/08/219,842  
FILING DATE: 30-MAR-1994  
CLASSIFICATION: 435  
ATTORNEY/AGENT INFORMATION:  
NAME: Campbell, Cathryn A.  
REGISTRATION NUMBER: 31,815  
REFERENCE/DOCKET NUMBER: P-AG 9504  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: (619) 535-9001  
TELEFAX: (619) 535-8949  
INFORMATION FOR SEQ ID NO: 3:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 856 base pairs  
TYPE: nucleic acid  
STRANDEDNESS: double  
TOPOLOGY: linear  
US-08-219-842-3

alignment\_scores:  
Quality: 64.00 Length: 28  
Ratio: 3.200 Gaps: 0  
Percent Similarity: 71.429 Percent Identity: 50.000

alignment\_block:  
US-09-712-768-8 x US-08-219-842-3 ..

Align seg 1/1 to: US-08-219-842-3 from: 1 to: 856

```
2 SerTlePheLeuLeuValCysLeuIleArgIleLeuArgGlyAlaMetse 18  
||| ||||| ::|||::||| ||| ::|||  
704 TCACTTTCCTCCTCATCTGCTTCATCCGCCAACATAATATTTCACCTTAC 753  
18 rAlaLysGlnHisValGlyPheGluMetAlaIa 29  
:::|||||::||| ||||| |||||  
754 ATCCAAACATCATCTTGGCTTGAAGCGCCGCC 787
```

seq\_name: /cgn2\_6/ptodata/2/1na/5A\_COMB.seq:US-08-451-096-3

seq\_documentation\_block:

Sequence 3, Application US/08451096  
Patent No. 5760205  
GENERAL INFORMATION:  
APPLICANT: Parker, W. D.  
APPLICANT: Herrnstadt, Corinna  
TITLE OF INVENTION: Diagnostic and Therapeutic Compositions  
NUMBER OF SEQUENCES: 95  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: Campbell and Flores  
STREET: 4370 La Jolla Village Drive, Suite 700  
CITY: San Diego  
STATE: California  
COUNTRY: USA  
ZIP: 92122  
COMPUTER READABLE FORM:  
MEDIUM TYPE: Floppy disk  
COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: Patentin Release #1.0, Version #1.25  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/08/451,096  
FILING DATE:  
CLASSIFICATION: 435  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: US 08/219,842  
FILING DATE: 30-MAR-1994

```

; ATTORNEY/AGENT INFORMATION:
; NAME: Campbell, Cathryn A.
; REGISTRATION NUMBER: 31,815
; REFERENCE/DOCKET NUMBER: P-AG 9504
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (619) 535-9001
; TELEFAX: (619) 535-8949
; INFORMATION FOR SEQ ID NO: 3:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 856 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: double
; TOPOLOGY: linear
US-08-451-096-3

alignment_scores:
  Quality: 64.00      Length: 28
  Ratio: 3.200        Gaps: 0
  Percent Similarity: 71.429   Percent Identity: 50.000

alignment_block:
US-09-712-768-8 x US-08-451-096-3 ..

Align seq 1/1 to: US-08-451-096-3 from: 1 to: 856

2 SerTlePheLeuValCysLeuIleArgIleLeuArgGlyAlaMetSe 18
||| ||||| ::|||::||| ||| ::|||
704 TCACCTTCTCCTACTATCTGCTTCATCCGCACTAATATTTCACCTTAC 753
18 rAlaLysGlnHisValGlyPheGluMetAla 29
::|||::||| ||||| |||||
754 ATCCAACATCATCTTGCTTGCGATTGGAAGCGCGCCGCC 787

seq_name: /cgn2_6/ptodata/2/ina/5B_COMB.seq:US-08-810-599-82

seq_documentation_block:
; Sequence 82, Application US/08810599
; Patent No. 5976798
; GENERAL INFORMATION:
; APPLICANT: PARKER, W. Davis
; APPLICANT: HERRNSTADT, Corinna
; APPLICANT: GHOSH, Soumitra S.
; APPLICANT: FAHY, Eoin
; TITLE OF INVENTION: Methods for Detecting Mitochondrial Mutations
; TITLE OF INVENTION: Diagnostic for Alzheimer's Disease and Methods for Determining
; TITLE OF INVENTION: of Mitochondrial Nucleic Acid
; NUMBER OF SEQUENCES: 82
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Kenyon & Kenyon
; STREET: 1025 Connecticut Avenue, N.W., Suite 600
; CITY: Washington
; STATE: D.C.
; COUNTRY: US
; ZIP: 20036
; COMPUTER READABLE FORM:
; MEDIUM TYPE: 3.25" floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Wordperfect 6.1 for Windows
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/810,599
; FILING DATE: Concurrent Herewith
; CLASSIFICATION: 436
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 08/757,438
; FILING DATE: 27 Nov. 5976798 1996
; APPLICATION NUMBER: US 08/614,072
; FILING DATE: 12 Mar 1996
; APPLICATION NUMBER: US 08/536,036
; FILING DATE: 29 Sep 1995
; APPLICATION NUMBER: US 08/414,969
; FILING DATE: 31 Mar 1995

```

```

; APPLICATION NUMBER: US 08/413,740
; FILING DATE: 30 Mar 1995
; APPLICATION NUMBER: US 08/410,658
; FILING DATE: 24 MARCH 1995
; APPLICATION NUMBER: US 08/397,808
; FILING DATE: 3 Mar 1995
; APPLICATION NUMBER: US 08/219,842
; FILING DATE: 30 MARCH 1994
; ATTORNEY/AGENT INFORMATION:
; NAME: Tofteneftl, Judith L.
; REGISTRATION NUMBER: 39,048
; REFERENCE/DOCKET NUMBER: 2105/17
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 202-429-1776
; TELEFAX: 202-429-0796
; INFORMATION FOR SEQ ID NO: 82:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 926 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: double
; TOPOLOGY: linear
; MOLECULE TYPE: other nucleic acid
; HYPOTHEICAL: NO
; ANTI-SENSE: NO
US-08-810-599-82

alignment_scores:
  Quality: 64.00      Length: 28
  Ratio: 3.200        Gaps: 0
  Percent Similarity: 71.429   Percent Identity: 50.000

alignment_block:
US-09-712-768-8 x US-08-810-599-82 ..

Align seq 1/1 to: US-08-810-599-82 from: 1 to: 926

2 SerTlePheLeuValCysLeuIleArgIleLeuArgGlyAlaMetSe 18
||| ||||| ::|||::||| ||| ::|||
704 TCACCTTCTCCTACTATCTGCTTCATCCGCACTAATATTTCACCTTAC 753
18 rAlaLysGlnHisValGlyPheGluMetAla 29
::|||::||| ||||| |||||
754 ATCCAACATCATCTTGCTTGCGATTGGAAGCGCGCCGCC 787

seq_name: /cgn2_6/ptodata/2/ina/6B_COMB.seq:US-08-413-740A-3

seq_documentation_block:
; Sequence 3, Application US/08413740A
; Patent No. 6171859
; GENERAL INFORMATION:
; APPLICANT: HERRNSTADT, CORINNA
; APPLICANT: PARKER, WILLIAM D.
; APPLICANT: DAVIS, ROBERT
; APPLICANT: MILLER, SCOTT W.
; TITLE OF INVENTION: Diagnosis, Therapy and Cellular and
; TITLE OF INVENTION: Animal Models for Diseases Associated with Mitochondrial
; NUMBER OF SEQUENCES: 206
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Kenyon & Kenyon
; STREET: 1025 Connecticut Avenue, N.W.
; CITY: Washington
; STATE: DC
; COUNTRY: USA
; ZIP: 20036-5405
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/413,740A

```

```

; FILING DATE:
; CLASSIFICATION: 435
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: PCT/US95/04063
; FILING DATE: 30-MAR-1995
; APPLICATION NUMBER: 08/413,740
; FILING DATE: 30-MAR-1995
; CLASSIFICATION: 435
; ATTORNEY/AGENT INFORMATION:
; NAME: Bonham, David B.
; REGISTRATION NUMBER: 34297
; REFERENCE/DOCKET NUMBER: 2105/7
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (202) 429-0796
; TELEFAX: (202) 429-0796
; INFORMATION FOR SEQ ID NO: 3:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 954 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: double
; TOPOLOGY: linear
; MOLECULE TYPE: other nucleic acid
; HYPOTHEICAL: NO
; ANTI-SENSE: NO
; US-08-413-740A-3

alignment_scores:
      Quality: 64.00      Length: 28
      Ratio: 3.200      Gaps: 0
      Percent Similarity: 71.429      Percent Identity: 50.000

alignment_block:
US-09-712-768-8 x US-08-413-740A-3 ..

Align seg 1/1 to: US-08-413-740A-3 from: 1 to: 954

2 SerTleheuleuValCysLeuileuArgileuArglyAlaMetse 18
||| ||||| ::|||::||| ||| ::|||
703 TCACTTCTCCTACTATCTGCTTCATCCGCCACATATATTCACCTTAC 752
18 rAlaLysGlnHisValGlyPheGluMetAlaIa 29
::|||::||| ||||| |||||
753 ATCCAAACATCATCTTGCTTGCATGCAAGCGCGCCG 786

seq_name: /cgn2_6/ptodata/2/ina/PCTUS_COMB.seq:PCT-US95-04063-3

seq_documentation_block:
; Sequence 3, Application PC/TUS9504063
; GENERAL INFORMATION:
; APPLICANT: HERRNSTADT, CORINNA
; APPLICANT: PARKER, WILLIAM D.
; APPLICANT: DAVIS, ROBERT W.
; APPLICANT: MILLER, SCOTT W.
; TITLE OF INVENTION: Diagnosis, Therapy and Cellular and
; TITLE OF INVENTION: Animal Models for Diseases Associated with Mitochondrial
; TITLE OF INVENTION: Defects
; NUMBER OF SEQUENCES: 206
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Kenyon & Kenyon
; STREET: 1025 Connecticut Avenue, N.W.
; CITY: Washington
; STATE: DC
; COUNTRY: USA
; ZIP: 20036-5405
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: PCT/US95/04063
; FILING DATE: 30-MAR-1995

```

```

; CLASSIFICATION:
; ATTORNEY/AGENT INFORMATION:
; NAME: Bonham, David B.
; REGISTRATION NUMBER: 34297
; REFERENCE/DOCKET NUMBER: 2105/7
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (202) 429-0796
; TELEFAX: (202) 429-0796
; INFORMATION FOR SEQ ID NO: 3:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 954 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: double
; TOPOLOGY: linear
; MOLECULE TYPE: other nucleic acid
; HYPOTHEICAL: NO
; ANTI-SENSE: NO
; PCT-US95-04063-3

alignment_scores:
      Quality: 64.00      Length: 28
      Ratio: 3.200      Gaps: 0
      Percent Similarity: 71.429      Percent Identity: 50.000

alignment_block:
US-09-712-768-8 x PCT-US95-04063-3 ..

Align seg 1/1 to: PCT-US95-04063-3 from: 1 to: 954

2 SerTleheuleuValCysLeuileuArgileuArglyAlaMetse 18
||| ||||| ::|||::||| ||| ::|||
703 TCACTTCTCCTACTATCTGCTTCATCCGCCACATATATTCACCTTAC 752
18 rAlaLysGlnHisValGlyPheGluMetAlaIa 29
::|||::||| ||||| |||||
753 ATCCAAACATCATCTTGCTTGCATGCAAGCGCGCCG 786

seq_name: /cgn2_6/ptodata/2/ina/6B_COMB.seq:US-09-097-889-2

seq_documentation_block:
; Sequence 2, Application US/09097889
; Patent No. 6218117
; GENERAL INFORMATION:
; APPLICANT: Herrnstadt, Corinna
; APPLICANT: Ghosh, Soumitra S.
; APPLICANT: Davis, Robert E.
; TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR IDENTIFYING
; TITLE OF INVENTION: AGENTS THAT QUANTITATIVELY ALTER DETECTABLE
; TITLE OF INVENTION: EXTRAMITOCHONDRIAL DNA: MITOCHONDRIAL DNA RATIOS
; NUMBER OF SEQUENCES: 26
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: SEED and BERRY LLP
; STREET: 6300 Columbia Center, 701 Fifth Avenue
; CITY: Seattle
; STATE: Washington
; COUNTRY: USA
; ZIP: 98104
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/09/097,889
; FILING DATE: 15-JUN-1998
; CLASSIFICATION: 435
; ATTORNEY/AGENT INFORMATION:
; NAME: Roseman Ph.D., Stephen J.
; REGISTRATION NUMBER: 43,058
; REFERENCE/DOCKET NUMBER: 660088.417
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (206) 622-4900

```



```
; TELEFAX: (206) 682-6031
; INFORMATION FOR SEQ ID NO: 2:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 16569 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; US-09-097-889-2

alignment_scores:
  Quality: 64.00      Length: 28
  Ratio: 3.200      Gaps: 0
  Percent Similarity: 71.429      Percent Identity: 50.000

alignment_block:
US-09-712-768-8 x US-09-097-889-2 ..

Align seg 1/1 to: US-09-097-889-2 from: 1 to: 16569

2 SerilepheuValCysleuileArgIleleuArgIylametse 18
||| ||||| ::|||::||| |||
9840 TCACACTTCTCTCACTATCTCTTCATCCGCCAACAATATTTCATTTC 9889
18 rAlaLysGlnHisValGlyPheGluMetAlaAla 29
::|||::||| ||||| |||||
9890 ATCCAAACATCACTTTCCTTCGACGCGCGCGC 9923

seq_name: /cgn2_6/ptodata/2/ina/5A_COMB.seq:US-08-920-812-8

seq_documentation_block:
; Sequence 8, Application US/08920812
; Patent No. 5763188
; GENERAL INFORMATION:
; APPLICANT: Ohno, Tsuneya
; APPLICANT: Matsuhisa, Akio
; APPLICANT: Uehara, Hirotsugu
; APPLICANT: Eda, Soji
; TITLE OF INVENTION: Probe for Diagnosing Infectious Disease
; NUMBER OF SEQUENCES: 25
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Marshall, O'Toole, Gerstein, Murray & Borun
; STREET: 6300 Sears Tower, 233 South Wacker Drive
; CITY: Chicago
; STATE: Illinois
; COUNTRY: United States of America
; ZIP: 60606-6402
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/920,812
; FILING DATE: 29-AUG-1997
; CLASSIFICATION: 435
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 08/362,577
; FILING DATE: 27-MAR-1995
; ATTORNEY/AGENT INFORMATION:
; NAME: Rin-Laures, Li-Hsien
; REGISTRATION NUMBER: 33,547
; REFERENCE/DOCKET NUMBER: 19036/32420
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 312/474-6300
; TELEFAX: 312/474-0448
; TELEX: 25-3856
; INFORMATION FOR SEQ ID NO: 8:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 3287 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: double
; TOPOLOGY: linear
```

```
; MOLECULE TYPE: Genomic DNA
; ORIGINAL SOURCE:
; ORGANISM: Staphylococcus epidermidis
; STRAIN: Clinical Isolate SE-37
; US-08-920-812-8

alignment_scores:
  Quality: 58.00      Length: 25
  Ratio: 2.900      Gaps: 0
  Percent Similarity: 80.000      Percent Identity: 40.000

alignment_block:
US-09-712-768-8 x US-08-920-812-8 ..

Align seg 1/1 to: US-08-920-812-8 from: 1 to: 3287

5 LeuLeuValCysleuileArgIleleuArgIylametSerAlaLysG1 21
||||| ||||| ::|||::||| |||
2607 CTACTGTTTGTCTAGTCAGATTATTAAGTCCCACTACTGCTGATAG 2656
21 nhisValGlyPheGluMetAlaAla 29
: |||::|||:: |||||
2657 AGTTGATCATTTGATGCGCTCGAGT 2681

seq_name: /cgn2_6/ptodata/2/ina/5A_COMB.seq:US-08-920-827-8

seq_documentation_block:
; Sequence 8, Application US/08920827
; Patent No. 5770375
; GENERAL INFORMATION:
; APPLICANT: Ohno, Tsuneya
; APPLICANT: Matsuhisa, Akio
; APPLICANT: Uehara, Hirotsugu
; APPLICANT: Eda, Soji
; TITLE OF INVENTION: Probe for Diagnosing Infectious Disease
; NUMBER OF SEQUENCES: 25
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Marshall, O'Toole, Gerstein, Murray & Borun
; STREET: 6300 Sears Tower, 233 South Wacker Drive
; CITY: Chicago
; STATE: Illinois
; COUNTRY: United States of America
; ZIP: 60606-6402
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/920,827
; FILING DATE: 29-AUG-1997
; CLASSIFICATION: 435
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 08/362,577
; FILING DATE: 27-MAR-1995
; ATTORNEY/AGENT INFORMATION:
; NAME: Rin-Laures, Li-Hsien
; REGISTRATION NUMBER: 33,547
; REFERENCE/DOCKET NUMBER: 19036/32420
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 312/474-6300
; TELEFAX: 312/474-0448
; TELEX: 25-3856
; INFORMATION FOR SEQ ID NO: 8:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 3287 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: double
; TOPOLOGY: linear
; MOLECULE TYPE: Genomic DNA
; ORIGINAL SOURCE:
; ORGANISM: Staphylococcus epidermidis
```

STRAIN: Clinical Isolate SE-37  
US-08-920-827-8

alignment\_scores:  
Quality: 58.00 Length: 25  
Ratio: 2.900 Gaps: 0  
Percent Similarity: 80.000 Percent Identity: 40.000

alignment\_block:  
US-09-712-768-8 x US-08-920-827-8 ..

Align seg 1/1 to: US-08-920-827-8 from: 1 to: 3287

5 LeuLeuValCysLeuIleArgIleLeuArgGlyAlaMetSerAlaLysG1 21  
|||||  
2607 CTACTGTTGTTGCTAGTACGATTAAATTAAGGTCACACTGCTGATAG 2656  
21 nhISValGlyPheGluMetAlaIa 29  
: |||  
2657 AGTTGATCATTTGATGCTCGAGT 2681

seq\_name: /cgn2\_6/prodata/2/lna/5A\_COMB.seq:US-08-921-177-8

seq\_documentation\_block:  
Sequence 8, Application US/08921177  
Patent No. 5798211  
GENERAL INFORMATION:  
APPLICANT: Ohno, Tsuneya  
APPLICANT: Matsuhisa, Akio  
APPLICANT: Uehara, Hirotsugu  
APPLICANT: Eda, Soji  
TITLE OF INVENTION: Probe for Diagnosing Infectious Disease  
NUMBER OF SEQUENCES: 25  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: Marshall, O'Toole, Gerstein, Murray & Borun  
STREET: 6300 Sears Tower, 233 South Wacker Drive  
CITY: Chicago  
STATE: Illinois  
COUNTRY: United States of America  
ZIP: 60606-6402  
COMPUTER READABLE FORM:  
MEDIUM TYPE: Floppy disk  
COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: Patent In Release #1.0, Version #1.25  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/08/921.177  
FILING DATE: 29-AUG-1997  
CLASSIFICATION: 435  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: US 08/362,577  
FILING DATE: 27-MAR-1995  
ATTORNEY/AGENT INFORMATION:  
NAME: Rin-Laures, Li-Hsien  
REGISTRATION NUMBER: 33,547  
REFERENCE/DOCKET NUMBER: 19036/32420  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: 312/474-6300  
TELEFAX: 312/474-0448  
TELEX: 25-3856  
INFORMATION FOR SEQ ID NO: 8:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 3287 base pairs  
TYPE: nucleic acid  
STRANDEDNESS: double  
TOPOLOGY: linear  
MOLECULE TYPE: Genomic DNA  
ORGANISM: Staphylococcus epidermidis  
STRAIN: Clinical Isolate SE-37  
US-08-921-177-8

alignment\_scores:  
Quality: 58.00 Length: 25  
Ratio: 2.900 Gaps: 0  
Percent Similarity: 80.000 Percent Identity: 40.000

alignment\_block:  
US-09-712-768-8 x US-08-921-177-8 ..

Align seg 1/1 to: US-08-921-177-8 from: 1 to: 3287

5 LeuLeuValCysLeuIleArgIleLeuArgGlyAlaMetSerAlaLysG1 21  
|||||  
2607 CTACTGTTGTTGCTAGTACGATTAAATTAAGGTCACACTGCTGATAG 2656  
21 nhISValGlyPheGluMetAlaIa 29  
: |||  
2657 AGTTGATCATTTGATGCTCGAGT 2681

seq\_name: /cgn2\_6/prodata/2/lna/5A\_COMB.seq:US-08-362-577C-8

seq\_documentation\_block:  
Sequence 8, Application US/08362577C  
Patent No. 5807673  
GENERAL INFORMATION:  
APPLICANT: Ohno, Tsuneya  
APPLICANT: Matsuhisa, Akio  
APPLICANT: Uehara, Hirotsugu  
APPLICANT: Eda, Soji  
TITLE OF INVENTION: Probe for Diagnosing Infectious Disease  
NUMBER OF SEQUENCES: 25  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: Marshall, O'Toole, Gerstein, Murray & Borun  
STREET: 6300 Sears Tower, 233 South Wacker Drive  
CITY: Chicago  
STATE: Illinois  
COUNTRY: United States of America  
ZIP: 60606-6402  
COMPUTER READABLE FORM:  
MEDIUM TYPE: Floppy disk  
COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: Patent In Release #1.0, Version #1.25  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/08/362,577C  
FILING DATE: 27-MAR-1995  
CLASSIFICATION: 536  
ATTORNEY/AGENT INFORMATION:  
NAME: Rin-Laures, Li-Hsien  
REGISTRATION NUMBER: 33,547  
REFERENCE/DOCKET NUMBER: 19036/32420  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: 312/474-6300  
TELEFAX: 312/474-0448  
TELEX: 25-3856  
INFORMATION FOR SEQ ID NO: 8:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 3287 base pairs  
TYPE: nucleic acid  
STRANDEDNESS: double  
TOPOLOGY: linear  
MOLECULE TYPE: Genomic DNA  
ORGANISM: Staphylococcus epidermidis  
STRAIN: Clinical Isolate SE-37  
US-08-362-577C-8

alignment\_scores:  
Quality: 58.00 Length: 25  
Ratio: 2.900 Gaps: 0  
Percent Similarity: 80.000 Percent Identity: 40.000

alignment\_block:  
US-09-712-768-8 x US-08-362-577C-8 ..

Align seg 1/1 to: US-08-362-577C-8 from: 1 to: 3287

```

5 LeuLeuValCysLeuIleAArgIleLeuAArgIAlaMetSerAlaIysG1 21
|||||
2607 CTACTGTGTTGCTACTAGATTAATTAAAGTCCACACTACTGCTGATAG 2656
21 nHISValGlyPheGluMetAlaIa 29
: |||:||||: :
2657 AGTTGATCATTTGATGCTCGAGT 2681

```

seq\_name: /cgn2\_6/ptodata/2/ina/5B\_COMB.seq:US-08-920-828-8

```

seq_documentation_block:
; Sequence 8, Application US/08920828
; Patent No. 5853998
; GENERAL INFORMATION:
; APPLICANT: Ohno, Tsuneya
; APPLICANT: Matsuhisa, Akio
; APPLICANT: Uehara, Hirotsugu
; APPLICANT: Ega, Soji
; TITLE OF INVENTION: Probe for Diagnosing Infectious Disease
; NUMBER OF SEQUENCES: 25
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Marshall, O'Toole, Gerstein, Murray & Borun
; STREET: 6300 Sears Tower, 233 South Wacker Drive
; CITY: Chicago
; STATE: Illinois
; COUNTRY: United States of America
; ZIP: 60606-6402
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/920,828
; FILING DATE: 29-AUG-1997
; CLASSIFICATION: 435
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 08/362,577
; FILING DATE: 27-MAR-1995
; ATTORNEY/AGENT INFORMATION:
; NAME: Rin-Laures, Li-Hsien
; REGISTRATION NUMBER: 33,547
; REFERENCE/DOCKET NUMBER: 19036/32420
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 312/474-6300
; TELEFAX: 312/474-0448
; TELEX: 25-3856
; INFORMATION FOR SEQ ID NO: 8:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 3287 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: double
; TOPOLOGY: linear
; MOLECULE TYPE: Genomic DNA
; ORIGINAL SOURCE:
; ORGANISM: staphylococcus epidermidis
; STRAIN: clinical isolate SE-37
US-08-920-828-8

```

## alignment\_scores:

Quality: 58.00 Length: 25  
Ratio: 2.900 Gaps: 0  
Percent Similarity: 80.000 Percent Identity: 40.000

alignment\_block:  
US-09-712-768-8 x US-08-920-828-8 ..

Align seg 1/1 to: US-08-920-828-8 from: 1 to: 3287

```

5 LeuLeuValCysLeuIleAArgIleLeuAArgIAlaMetSerAlaIysG1 21
|||||
2607 CTACTGTGTTGCTACTAGATTAATTAAAGTCCACACTACTGCTGATAG 2656
21 nHISValGlyPheGluMetAlaIa 29
: |||:||||: :
2657 AGTTGATCATTTGATGCTCGAGT 2681

```

seq\_name: /cgn2\_6/ptodata/2/ina/5A\_COMB.seq:US-08-484-105-7

```

seq_documentation_block:
; Sequence 7, Application US/08484105.
; Patent No. 589341
; GENERAL INFORMATION:
; APPLICANT: STILIMAN, Bruce
; APPLICANT: BELL, Stephen P
; APPLICANT: KOBAYASHI, Ryuji
; APPLICANT: KINE, Jasper
; APPLICANT: FOSS, Margit
; APPLICANT: McNALLY, Francis J
; APPLICANT: LAURENSEN, Patricia
; APPLICANT: HERSKOWITZ, Ira
; APPLICANT: LI, Joachim J
; APPLICANT: GAVIN, Kimberly
; TITLE OF INVENTION: ORIGIN OF REPLICATION COMPLEX GENES
; NUMBER OF SEQUENCES: 24
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: FLEHR, HOHBACH, TEST, ALBRITTON & HERBERT
; STREET: 4 Embarcadero Center, Suite 3400
; CITY: San Francisco
; STATE: California
; COUNTRY: USA
; ZIP: 94111-4187
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/484,105
; FILING DATE:
; CLASSIFICATION: 435
; ATTORNEY/AGENT INFORMATION:
; NAME: Osman Ph.D., Richard Aron
; REGISTRATION NUMBER: 36,627
; REFERENCE/DOCKET NUMBER: A-59032/DJB/RAO
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (415) 494-8700
; TELEFAX: (415) 494-8771
; TELEX: 910 277299
; INFORMATION FOR SEQ ID NO: 7:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 2404 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: double
; TOPOLOGY: linear
; MOLECULE TYPE: cDNA
US-08-484-105-7

```

## alignment\_scores:

Quality: 49.00 Length: 24  
Ratio: 2.450 Gaps: 0  
Percent Similarity: 83.333 Percent Identity: 41.667

alignment\_block:  
US-09-712-768-8 x US-08-484-105-7/rev ..

Align seg 1/1 to reverse of: US-08-484-105-7 from: 1 to: 2404

2 SerIlePheLeuLeuValCysLeuIleAArgIleLeuAArgIAlaMetSerAlaIysG1 18



GenCore version 4.5  
Copyright (c) 1993 - 2000 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: January 10, 2002, 02:00:24 ; Search time 18.96 Seconds

(without alignments)  
116,512 Million cell updates/sec

Title: US-09-712-768-8

Perfect score: 142

Sequence: 1 DSIFLVCLIRILRGAMSAKQHVGFEMAA 29

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 219241 seqs, 76174552 residues

Total number of hits satisfying chosen parameters: 219241

Minimum DB seq length: 0  
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%  
Listing first 45 summaries

Database :

1: p1r1:\*  
2: p1r2:\*  
3: p1r3:\*  
4: p1r4:\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

#### SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	105	73.9	274	2	S03807
2	89	62.7	266	2	C45164
3	79	55.6	281	2	T11235
4	78	54.9	261	2	T11330
5	78	54.9	265	1	OTWT3M
6	77	54.2	272	2	F58930
7	75	52.8	261	1	OTMS3
8	75	52.8	261	2	S55010
9	75	52.8	261	2	S04753
10	74	52.1	265	1	OTR23M
11	74	52.1	265	2	S20801
12	73	51.4	260	2	T11798
13	73	51.4	261	2	T11199
14	73	51.4	265	2	A48304
15	71	50.0	262	1	OTFF3
16	71	50.0	262	1	OTFF3Y
17	70	49.3	261	2	T11186
18	70	49.3	261	2	T11146
19	70	49.3	261	2	G59153
20	69	48.6	261	2	B40076
21	69	48.6	261	2	T11092
22	69	48.6	261	2	T11826
23	69	48.6	261	2	T11839
24	69	48.6	261	2	T11028
25	69	48.6	261	2	T11469
26	69	48.6	265	1	OTSY3M
27	69	48.6	272	2	S43951
28	68	47.9	260	2	H34284
29	68	47.9	261	2	T11525

#### ALIGNMENTS

30	68	47.9	261	2	T11540	cytochrome-c oxida
31	68	47.9	261	2	T11107	cytochrome-c oxida
32	68	47.9	265	2	S25951	cytochrome-c oxida
33	68	47.9	272	2	S57460	cytochrome-c oxida
34	68	47.9	295	2	A83632	cytochrome-c oxida
35	68	47.9	311	2	S66600	cytochrome-c oxida
36	67	47.2	260	1	OTXL3	cytochrome-c oxida
37	67	47.2	260	2	T11133	cytochrome-c oxida
38	67	47.2	261	2	S10193	cytochrome-c oxida
39	67	47.2	261	2	S26157	cytochrome-c oxida
40	67	47.2	261	2	T10978	cytochrome-c oxida
41	67	47.2	261	2	T11173	cytochrome-c oxida
42	67	47.2	262	2	T11922	cytochrome-c oxida
43	66	46.5	261	1	OTR03	cytochrome-c oxida
44	66	46.5	261	2	T11512	cytochrome-c oxida
45	66	46.5	261	2	T11293	cytochrome-c oxida

RESULT 1  
S03807

Cytochrome-c oxidase (EC 1.9.3.1) chain III - Paracoccus denitrificans  
C:Species: Paracoccus denitrificans  
C:Date: 28-Feb-1990 #sequence\_revision 28-Feb-1990 #text\_change 11-Jun-1999  
C:Accession: S03807; S02524; A24371  
R:Raitio, M.; Jalli, T.; Saraste, M.  
EMBO J. 6, 2825-2833, 1987

A:Title: The Paracoccus denitrificans cytochrome aa(3) has a third subunit.  
A:Reference number: S03803  
A:Accession: S03807

A:Molecule type: DNA  
A:Residues: 1-274 <RAI>  
A:Cross-references: EMBL:X05828; NID:945468; PIDN:CAA29272.1; PID:945473

R:Hall, T.; Puustinen, A.; Finel, M.  
Eur. J. Biochem. 172, 543-546, 1988

A:Title: The Paracoccus denitrificans cytochrome aa(3) has a third subunit.  
A:Reference number: S02524; MUID:88166732

A:Accession: S02524  
A:Molecule type: protein

A:Residues: 2-16 <HAL>  
R:Saraste, M.; Raitio, M.; Peramaa, A.

FEBS Lett. 206, 154-156, 1986  
A:Title: A gene in Paracoccus for subunit III of cytochrome oxidase.  
A:Reference number: A24371; MUID:87005242

A:Accession: A24371  
A:Molecule type: DNA  
A:Residues: 1-144 <SAR>

A:Cross-references: GB:X04406; NID:945466; PIDN:CAA27995.1; PID:945467  
C:Genetics:  
A:Gene: COIII

C:Superfamily: cytochrome-c oxidase chain III  
C:Keywords: electron transfer; membrane-associated complex; oxidative phosphorylation  
F:2-274/Product: cytochrome-c oxidase chain III #status experimental <MAN>

Query Match 73.9%; Score 105; DB 2; Length 274;  
Best Local Similarity 71.4%; Pred. No. 1.3e-08;

Matches 20; Conservative 4; Mismatches 4; Indels 0; Gaps 0;

QY 2 DSIFLVCLIRILRGAMSAKQHVGFEMAA 29  
DB 225 TIFLVCLIRILRGAMSAKQHVGFEMAA 252

RESULT 2  
C45164

Cytochrome-c oxidase (EC 1.9.3.1) chain III - Rhodobacter sphaeroides  
C:Species: Rhodobacter sphaeroides  
C:Date: 26-May-1994 #sequence\_revision 26-May-1994 #text\_change 17-Mar-1999  
C:Accession: C45164  
R:cao, J.; Hostler, J.; Shapleigh, J.; Revzin, A.; Ferguson-Miller, S.





A:Status: preliminary  
 A:Molecule type: DNA  
 A:Residues: 1-272 <ARN>  
 A:Cross-references: GB:D89861; NID:94115761; PIDN:BAA34657.1; PID:93927869  
 C:Genetics:  
 A:Gene: cox3  
 A:Genome: mitochondrion  
 C:Superfamily: cytochrome-c oxidase chain III  
 C:Keywords: electron transfer; membrane-associated complex; mitochondrial inner membrane

Query Match 54.2%; Score 77; DB 2; Length 272;  
 Best Local Similarity 50.0%; Pred. NO. 0.00027;  
 Matches 14; Conservative 7; Mismatches 7; Indels 0; Gaps 0;

QY 2 SIFLVCLIRIRGAMSAKOHVFEMAA 29  
 DB 222 SIFLVCLIRIRGAMSAKOHVFEMAA 249

RESULT 7  
 OTMS3

Cytochrome-c oxidase (EC 1.9.3.1) chain III - mouse mitochondrion  
 C:Species: mitochondrion Mus musculus (house mouse)  
 C:Date: 02-Apr-1982 #sequence\_revision 02-Apr-1982 #text\_change 07-Dec-1999  
 C:Accession: A00484  
 R:Bibb, M.J.; Van Etten, R.A.; Wright, C.T.; Walberg, M.W.; Clayton, D.A.  
 Cell 26, 167-180, 1981  
 A:Title: Sequence and gene organization of mouse mitochondrial DNA.  
 A:Reference number: A00153; MUID:82137051  
 A:Accession: A00484  
 A:Molecule type: DNA  
 A:Residues: 1-261 <BIB>  
 A:Cross-references: GB:J01420; NID:9342520; PIDN:AAB48650.1; PID:9896297; EMBL:V00711; N  
 C:Genetics:  
 A:Gene: coxi1  
 A:Genome: mitochondrion  
 A:Genetic code: SGC1  
 C:Superfamily: cytochrome-c oxidase chain III  
 C:Keywords: electron transfer; membrane-associated complex; mitochondrial inner membrane  
 F:16-34/Domain: transmembrane helix #status predicted <TR01>  
 F:41-66/Domain: transmembrane helix #status predicted <TR02>  
 F:73-105/Domain: transmembrane helix #status predicted <TR03>  
 F:129-155/Domain: transmembrane helix #status predicted <TR04>  
 F:156-183/Domain: transmembrane helix #status predicted <TR05>  
 F:191-223/Domain: transmembrane helix #status predicted <TR06>  
 F:233-256/Domain: transmembrane helix #status predicted <TR07>

Query Match 52.8%; Score 75; DB 1; Length 261;  
 Best Local Similarity 53.6%; Pred. NO. 0.00053;  
 Matches 15; Conservative 5; Mismatches 8; Indels 0; Gaps 0;

QY 2 SIFLVCLIRIRGAMSAKOHVFEMAA 29  
 DB 212 SIFLVCLIRIRGAMSAKOHVFEMAA 239

RESULT 8  
 S55010

Cytochrome-c oxidase (EC 1.9.3.1) chain III - sea lamprey mitochondrion  
 C:Species: mitochondrion Petromyzon marinus (sea lamprey)  
 C:Date: 23-Aug-1995 #sequence\_revision 03-Nov-1995 #text\_change 07-Dec-1999  
 C:Accession: S55010  
 R:Lee, W.J.; Kocher, J.D.  
 Genetics 139, 873-887, 1995  
 A:Title: Complete sequence of a sea lamprey (Petromyzon marinus) mitochondrial genome: e  
 A:Reference number: S54999; MUID:95229067  
 A:Accession: S55010  
 A:Molecule type: DNA  
 A:Residues: 1-261 <LEE>  
 A:Cross-references: EMBL:U11880; NID:9515484; PIDN:AAB08744.1; PID:9515492  
 C:Genetics:

A:Gene: coIII  
 A:Genome: mitochondrion  
 A:Genetic code: SGC1  
 C:Superfamily: cytochrome-c oxidase chain III  
 C:Keywords: electron transfer; membrane-associated complex; mitochondrial inner membr

Query Match 52.8%; Score 75; DB 2; Length 261;  
 Best Local Similarity 53.6%; Pred. NO. 0.00053;  
 Matches 15; Conservative 5; Mismatches 8; Indels 0; Gaps 0;

QY 2 SIFLVCLIRIRGAMSAKOHVFEMAA 29  
 DB 212 SIFLVCLIRIRGAMSAKOHVFEMAA 239

RESULT 9  
 S04753

Cytochrome-c oxidase (EC 1.9.3.1) chain III - rat mitochondrion  
 C:Species: mitochondrion Rattus norvegicus (Norway rat)  
 C:Date: 07-Sep-1990 #sequence\_revision 07-Sep-1990 #text\_change 07-Dec-1999  
 C:Accession: S04753; S65372  
 R:Gadaleta, G.; Pepe, G.; De Candia, G.; Quagliariello, C.; Sbisa, E.; Saccone, C.  
 J. Mol. Evol. 28, 497-516, 1989  
 A:Title: The complete nucleotide sequence of the Rattus norvegicus mitochondrial geno  
 A:Reference number: S04747; MUID:89362487  
 A:Accession: S04753  
 A:Molecule type: DNA  
 A:Residues: 1-261 <GAD>  
 A:Cross-references: EMBL:X14848; NID:9854269; PIDN:CAA32960.1; PID:9639981  
 R:Schaegger, H.; Noack, H.; Halangk, W.; Brandt, U.; von Jagow, G.  
 Eur. J. Biochem. 230, 235-241, 1995  
 A:Title: Cytochrome-c oxidase in developing rat heart. Enzymic properties and amino-t  
 A:Reference number: S65372; MUID:95324529  
 A:Accession: S65372  
 A:Status: preliminary  
 A:Molecule type: protein  
 A:Residues: 75, 3-11 <SCH>  
 C:Genetics:  
 A:Genome: mitochondrion  
 A:Genetic code: SGC1  
 C:Superfamily: cytochrome-c oxidase chain III  
 C:Keywords: electron transfer; membrane-associated complex; mitochondrial inner membr

Query Match 52.8%; Score 75; DB 2; Length 261;  
 Best Local Similarity 53.6%; Pred. NO. 0.00053;  
 Matches 15; Conservative 5; Mismatches 8; Indels 0; Gaps 0;

QY 2 SIFLVCLIRIRGAMSAKOHVFEMAA 29  
 DB 212 SIFLVCLIRIRGAMSAKOHVFEMAA 239

RESULT 10  
 OTR23M

Cytochrome-c oxidase (EC 1.9.3.1) chain III - rice mitochondrion  
 C:Species: mitochondrion Oryza sativa (rice)  
 C:Date: 30-Jun-1992 #sequence\_revision 30-Jun-1992 #text\_change 20-Aug-1999  
 C:Accession: J00166; S29748; S07875  
 R:Kaleikau, E.K.; Andre, C.P.; Walbot, V.  
 Nucleic Acids Res. 18, 371, 1990  
 A:Title: Sequence of the rice mitochondrial gene for cytochrome oxidase subunit 3.  
 A:Reference number: J00166; MUID:90221829  
 A:Accession: J00166  
 A:Molecule type: DNA  
 A:Residues: 1-265 <KAL>  
 A:Cross-references: EMBL:X17040; NID:913216; PIDN:CAA34898.1; PID:913217  
 R:Li, A.W.; Narayanan, K.K.; Andre, C.P.; Kaleikau, E.K.; Walbot, V.  
 Curr. Genet. 21, 507-513, 1992  
 A:Title: Co-transcription of orf25 and coxIII in rice mitochondria.  
 A:Reference number: S26876; MUID:92315351  
 A:Accession: S29748

A:Status: translation not shown  
 A:Molecule type: DNA  
 A:Residues: 1-265 <LTIU>  
 A:Cross-references: EMBL:M74241; NID:g343214; PIDN:AAA66048.1; PID:g804815  
 C:Genetics:  
 A:Gene: cox3  
 A:Genome: mitochondrion  
 C:Superfamily: cytochrome-c oxidase chain III  
 C:Keywords: electron transfer; membrane-associated complex; mitochondrial inner membrane  
 F:1-35/Domain: transmembrane helix #status predicted <TR01>  
 F:44-69/Domain: transmembrane helix #status predicted <TR02>  
 F:76-108/Domain: transmembrane helix #status predicted <TR03>  
 F:132-155/Domain: transmembrane helix #status predicted <TR04>  
 F:159-186/Domain: transmembrane helix #status predicted <TR05>  
 F:194-226/Domain: transmembrane helix #status predicted <TR06>  
 F:236-259/Domain: transmembrane helix #status predicted <TR07>

Query Match 52.1%; Score 74; DB 1; Length 265;  
 Best Local Similarity 53.6%; Pred. No. 0.00076;  
 Matches 15; Conservative 5; Mismatches 8; Indels 0; Gaps 0;

QY 2 SIFLVCLIRIRGAMSAKOHVFEMAA 29  
 ::::||||| |::| ||||| ||  
 Db 215 TLFIVCGIRGYLGHITRKHHVGFEMAA 242

RESULT 11  
 cytochrome-c oxidase (EC 1.9.3.1) chain III - maize mitochondrion  
 C:Species: mitochondrion Zea mays (maize)  
 C:Date: 22-Nov-1993 #sequence\_revision 26-May-1995 #text\_change 20-Aug-1999  
 C:Accession: S20801  
 R:Kalelkar, E.K.; Malhot, V.  
 submitted to the EMBL Data Library, May 1990  
 A:Description: Allelic sequence of the maize mitochondrial gene for cytochrome oxidase  
 A:Reference number: S20801  
 A:Accession: S20801  
 A:Molecule type: DNA  
 A:Residues: 1-265 <KAL>  
 A:Cross-references: EMBL:X53055; NID:g14295; PIDN:CAA37222.1; PID:g14296  
 A:Experimental source: strain B37N; seedling  
 C:Genetics:  
 A:Gene: coxIII  
 A:Genome: mitochondrion  
 C:Superfamily: cytochrome-c oxidase chain III  
 C:Keywords: electron transfer; membrane-associated complex; mitochondrial inner membrane  
 F:2-16/Domain: mitochondrial matrix #status predicted <TM01>  
 F:17-35/Domain: transmembrane #status predicted <TM02>  
 F:36-43/Domain: intracristal #status predicted <TM03>  
 F:44-69/Domain: transmembrane #status predicted <TM02>  
 F:70-75/Domain: transmembrane #status predicted <TM02>  
 F:76-108/Domain: transmembrane #status predicted <TM03>  
 F:109-131/Domain: intracristal #status predicted <ITC2>  
 F:132-155/Domain: transmembrane #status predicted <TM04>  
 F:156-158/Domain: mitochondrial matrix #status predicted <MM3>  
 F:159-186/Domain: transmembrane #status predicted <TM05>  
 F:187-193/Domain: intracristal #status predicted <ITC3>  
 F:194-226/Domain: transmembrane #status predicted <TM06>  
 F:227-235/Domain: mitochondrial matrix #status predicted <MM4>  
 F:236-259/Domain: transmembrane #status predicted <TM07>  
 F:260-264/Domain: intracristal #status predicted <ITC4>

Query Match 52.1%; Score 74; DB 2; Length 265;  
 Best Local Similarity 53.6%; Pred. No. 0.00076;  
 Matches 15; Conservative 5; Mismatches 8; Indels 0; Gaps 0;

QY 2 SIFLVCLIRIRGAMSAKOHVFEMAA 29  
 ::::||||| |::| ||||| ||  
 Db 215 TLFIVCGIRGYLGHITRKHHVGFEMAA 242

RESULT 12  
 cytochrome-c oxidase (EC 1.9.3.1) chain III - sea urchin (Arbacia lixula) mitochondri  
 C:Species: mitochondrion Arbacia lixula (black urchin)  
 C:Date: 16-Jul-1999 #sequence\_revision 16-Jul-1999 #text\_change 21-Jul-2000  
 C:Accession: T11798; D39746; PH0855  
 R:De Giorgi, C.; Martiradonna, A.; Lanave, C.; Saccone, C.  
 Mol. Phylogenet. Evol. 5, 323-332, 1996  
 A:Title: Complete sequence of the mitochondrial DNA in the sea urchin Arbacia lixula:  
 A:Reference number: Z17346; MUID:96292527  
 A:Accession: T11798  
 A:Status: preliminary; translated from GB/EMBL/DBJ  
 A:Molecule type: DNA  
 A:Residues: 1-260 <DEG1>  
 A:Cross-references: EMBL:X80396; NID:g1321876; PIDN:CAA56613.1  
 A:Experimental source: eggs  
 R:De Giorgi, C.; Lanave, C.; Musci, M.D.; Saccone, C.  
 Mol. Biol. Evol. 8, 515-529, 1991  
 A:Title: Mitochondrial DNA in the sea urchin Arbacia lixula: evolutionary inferences  
 A:Reference number: A39746; MUID:92017217  
 A:Accession: D39746  
 A:Status: preliminary  
 A:Molecule type: DNA  
 A:Residues: 152-208, 'NS', 211-260 <DEG2>  
 A:Cross-references: EMBL:X53727; NID:g36368; PIDN:AAA98044.1; PID:g1280181  
 R:De Giorgi, C.; De Luca, F.; Saccone, C.  
 Gene 103, 249-252, 1991  
 A:Title: Mitochondrial DNA in the sea urchin Arbacia lixula: nucleotide sequence diff  
 A:Reference number: PH0854; MUID:91365253  
 A:Accession: PH0855  
 A:Molecule type: DNA  
 A:Residues: 157-210, 'H' <DEG3>  
 A:Cross-references: GB:M79455; GB:M35442; NID:g336366; PIDN:AAA31638.2; PID:g7304846  
 C:Genetics:  
 A:Genome: mitochondrion  
 A:Genetic code: SGC8  
 C:Superfamily: cytochrome-c oxidase chain III  
 C:Keywords: electron transfer; membrane-associated complex; mitochondrial inner membr

Query Match 51.4%; Score 73; DB 2; Length 260;  
 Best Local Similarity 57.7%; Pred. No. 0.0011;  
 Matches 15; Conservative 2; Mismatches 9; Indels 0; Gaps 0;

QY 4 FFLVCLIRIRGAMSAKOHVFEMAA 29  
 ||||| |::| || ||||| ||  
 Db 213 FFLVCLIRIRGAMSAKOHVFEMAA 238

RESULT 13  
 T11199  
 cytochrome-c oxidase (EC 1.9.3.1) chain III - Vidua chalybeata mitochondrion  
 C:Species: mitochondrion Vidua chalybeata  
 C:Date: 16-Jul-1999 #sequence\_revision 16-Jul-1999 #text\_change 07-Dec-1999  
 C:Accession: T11199  
 R:Mindell, D.P.; Sorenson, M.D.; Dimcheff, D.E.  
 Proc. Natl. Acad. Sci. U.S.A. 95, 10693-10697, 1998  
 A:Title: Multiple independent origins of mitochondrial gene order in birds.  
 A:Reference number: Z17242  
 A:Accession: T11199  
 A:Status: preliminary; translated from GB/EMBL/DBJ  
 A:Molecule type: DNA  
 A:Residues: 1-261 <MIN>  
 A:Cross-references: EMBL:AF090341; NID:g4894501; PID:g4894508; PIDN:AD032531.1  
 C:Genetics:  
 A:Genome: mitochondrion  
 A:Genetic code: SGC1  
 C:Superfamily: cytochrome-c oxidase chain III  
 C:Keywords: electron transfer; membrane-associated complex; mitochondrial inner membr

Query Match 51.4%; Score 73; DB 2; Length 261;  
 Best Local Similarity 50.0%; Pred. No. 0.0011;

Matches 14; Conservative 5; Mismatches 9; Indels 0; Gaps 0;

QY 2 SIFLVCLIRIRLGAMSAKOHVGFEMAA 29  
 |||||:|:| | ||| ||  
 Db 212 STFLVCLIRLRILKIVHFFPSNHFGEFAAA 239

## RESULT 14

A48304  
 cytochrome-c oxidase (EC 1.9.3.1) chain III - fava bean mitochondrion  
 C:Species: mitochondrion Vicia faba (fava bean)  
 C:Date: 31-Dec-1993 #sequence\_revision 31-Dec-1993 #text\_change 20-Aug-1999  
 C:Accession: A48304  
 R:Macfarlane, J.L.; Wahlertner, J.A.; Wolstenholme, D.R.  
 Curr. Genet. 17, 33-40, 1990  
 A:Title: A gene for cytochrome c oxidase subunit III (COXIII) in broad bean mitochondria  
 A:Reference number: A48304; MUID:90182727  
 A:Accession: A48304  
 A:Status: preliminary  
 A:Molecule type: DNA  
 A:Residues: 1-265 <MAC>  
 A:Cross-references: GB:X51690  
 C:Genetics:  
 A:Genome: mitochondrion  
 C:Superfamily: cytochrome-c oxidase chain III  
 C:Keywords: electron transfer; membrane-associated complex; mitochondrial inner membrane

Query Match 51.4%; Score 73; DB 2; Length 265;  
 Best Local Similarity 50.0%; Pred. No. 0.0011;  
 Matches 14; Conservative 6; Mismatches 8; Indels 0; Gaps 0;

QY 2 SIFLVCLIRIRLGAMSAKOHVGFEMAA 29  
 :|||:| | | | :|:| |||| ||  
 Db 215 TFLVLCIGIRQYLQGMTRKHVGFEMAA 242

## RESULT 15

## OTFF3

cytochrome-c oxidase (EC 1.9.3.1) chain III - fruit fly (Drosophila melanogaster) mitoch  
 C:Species: mitochondrion Drosophila melanogaster  
 C:Date: 18-Apr-1984 #sequence\_revision 18-Apr-1984 #text\_change 07-Dec-1999  
 C:Accession: A00485; A93463; S02250  
 R:de Bruijn, M.H.L.  
 Nature 304, 234-241, 1983  
 A:Title: Drosophila melanogaster mitochondrial DNA, a novel organization and genetic cod  
 A:Reference number: A93307; MUID:83245048  
 A:Accession: A00485  
 A:Molecule type: DNA  
 A:Residues: 1-179 <DEB>  
 R:Clary, D.O.; Wahlertner, J.A.; Wolstenholme, D.R.  
 Nucleic Acids Res. 11, 2411-2425, 1983  
 A:Title: Transfer RNA genes in Drosophila mitochondrial DNA: related 5' flanking sequenc  
 A:Reference number: A93463; MUID:83220794  
 A:Accession: A93463  
 A:Molecule type: DNA  
 A:Residues: 179-262 <CLA>  
 R:Garesse, R.  
 Genetics 118, 649-663, 1988  
 A:Title: Drosophila melanogaster mitochondrial DNA: gene organization and evolutionary c  
 A:Reference number: S01185; MUID:86212147  
 A:Accession: S02250  
 A:Molecule type: DNA  
 A:Residues: 179-262 <GAR>  
 A:Cross-references: GB:M37275; EMBL:Y00610; NID:g336819; PIDN:AAA69708.1; PID:g336820  
 C:Genetics:  
 A:Gene: COI11  
 A:Cross-references: FLYBase:FBgn0013676  
 A:Genome: mitochondrion  
 A:Genetic code: SGC4  
 C:Superfamily: cytochrome-c oxidase chain III  
 C:Keywords: electron transfer; membrane-associated complex; mitochondrial inner membrane  
 F:17-35/Domain: transmembrane helix #status predicted <TR01>

F:42-67/Domain: transmembrane helix #status predicted <TR02>  
 F:74-106/Domain: transmembrane helix #status predicted <TR03>  
 F:130-153/Domain: transmembrane helix #status predicted <TR04>  
 F:157-184/Domain: transmembrane helix #status predicted <TR05>  
 F:192-224/Domain: transmembrane helix #status predicted <TR06>  
 F:234-257/Domain: transmembrane helix #status predicted <TR07>

Query Match 50.0%; Score 71; DB 1; Length 262;  
 Best Local Similarity 57.7%; Pred. No. 0.0022;  
 Matches 15; Conservative 1; Mismatches 10; Indels 0; Gaps 0;

QY 4 FLVCLIRIRIRLGAMSAKOHVGFEMAA 29  
 |||||:| | | | | ||| ||  
 Db 215 FLVCLIRLRILNHNHPSKNHFGFEFAA 240

Search completed: January 10, 2002, 02:00:24  
 Job time: 547 sec



GenCore version 4.5  
Copyright (c) 1993 - 2000 Compugen Ltd.

OM protein - protein search, using sw model

Run on: January 10, 2002, 02:03:12 ; Search time 16.73 Seconds

(without alignments)  
63.555 Million cell updates/sec

Title: US-09-712-768-8

Perfect score: 142

Sequence: 1 DSIFLLVCLIRILGAMSAKQHVGFEMAA 29

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 100059 seqs, 36664827 residues

Total number of hits satisfying chosen parameters: 100059

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : SwissProt.39.\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

## SUMMARIES

Result No.	Score	Query	Length	DB ID	Description
1	105	73.9	273	1 COX3_PARDE	P06030 paracoccus
2	78	54.9	265	1 COX3_AEGCO	O36952 aegilops co
3	78	54.9	265	1 COX3_WHEAT	P15953 triticum ae
4	77	54.2	147	1 COX3_SPOFR	O35826 spodoptera
5	75	52.8	259	1 COX3_LOBL	O47475 loligo blec
6	75	52.8	260	1 COX3_RAT	P05505 rattus norv
7	75	52.8	261	1 COX3_MOUSE	P00416 mus muscul
8	75	52.8	261	1 COX3_PETMA	O35539 petromyzon
9	74	52.1	265	1 COX3_MAIZE	P09138 zea mays (m
10	74	52.1	265	1 COX3_ORYSA	P14852 oryza sativ
11	73	51.4	109	1 COX3_ARBLI	O33752 arabidopsi
12	73	51.4	265	1 COX3_VICFA	O03327 vicia faba
13	71	50.0	262	1 COX3_DROME	P00417 drosophila
14	71	50.0	262	1 COX3_DROYA	P00418 drosophila
15	69	48.6	261	1 COX3_GADMO	O95707 gadus morhu
16	69	48.6	261	1 COX3_HYLLA	O95707 hyllobates 1
17	69	48.6	263	1 COX3_LOCM1	P92514 arabidopsi
18	69	48.6	265	1 COX3_ARATH	P14853 glycine max
19	69	48.6	265	1 COX3_SOYBN	P14853 glycine max
20	69	48.6	272	1 COX3_CHOCR	P48872 chondrus cr
21	68	47.9	261	1 COX3_PARLI	P12702 paracentrot
22	68	47.9	261	1 COX3_PELSU	O79602 pelomedusa
23	68	47.9	261	1 COX3_SOUAC	O92448 squallus aca
24	68	47.9	261	1 COX3_STRCA	P26140 struthio ca
25	68	47.9	265	1 COX3_MARPO	P26888 marcanthia
26	68	47.9	265	1 COX3_MYED	P41775 mytilus edu
27	68	47.9	272	1 COX3_PYLTI	O37600 pylaia 1
28	67	47.2	260	1 COX3_XENLA	P00419 xenopus lae
29	67	47.2	261	1 COX3_CHICK	P18945 gallus gall
30	67	47.2	261	1 COX3_PHOVI	O00529 phoca vitul
31	67	47.2	261	1 COX3_PIG	O35916 sus scrofa
32	67	47.2	262	1 COX3_PROMI	O37600 protoltheca
33	66	46.3	261	1 COX3_CERSI	O03201 ceratotheri

## ALIGNMENTS

RESULT 1	ID	COX3_PARDE	STANDARD:	PRT:	273 AA.
AC	P06030,				
DT	13-AUG-1987 (Rel. 05, Created)				
DT	01-NOV-1990 (Rel. 16, Last sequence update)				
DT	20-AUG-2001 (Rel. 40, Last annotation update)				
DE	CYTOCHROME C OXIDASE POLYPEPTIDE III (EC 1.9.3.1) (CYTOCHROME AA3				
DE	SUBUNIT 3) (OXIDASE AA(3) SUBUNIT 3).				
GN	CTAE OR COIIT.				
OS	Paracoccus denitrificans.				
OC	Bacteria; Proteobacteria; alpha subdivision; Rhodobacter group;				
OC	Paracoccus.				
OX	NCBI_TaxID=266;				
RN	[1]				
RP	SEQUENCE FROM N.A.				
RC	STRAIN=SI657;				
RA	Raitio M., Jalli T., Saraste M.;				
RT	"Isolation and analysis of the genes for cytochrome c oxidase in				
RT	Paracoccus denitrificans.";				
RL	EMBO J. 6:2825-2833(1987).				
RN	[2]				
RP	SEQUENCE OF 1-143 FROM N.A.				
RC	STRAIN=SI657;				
RX	MEDLINE=87005242; PubMed=3019767;				
RA	Saraste M., Raitio M., Jalli T., Peramaa A.;				
RT	"A gene in Paracoccus for subunit III of cytochrome oxidase.";				
RL	FEBS Lett. 206:154-156(1986).				
RN	[3]				
RP	SEQUENCE OF 1-15.				
RA	MEDLINE=8816732; PubMed=2832167;				
RT	Haltia T., Puustinen A., Finel M.;				
RT	"The Paracoccus denitrificans cytochrome a3 has a third subunit.";				
RL	Eur. J. Biochem. 172:543-546(1988).				
RN	[4]				
RP	X-RAY CRYSTALLOGRAPHY (2.8 ANGSTROMS).				
RX	MEDLINE=95379947; PubMed=7651515;				
RA	Iwata S., Ostermeier C., Ludwig B., Michel H.;				
RT	"Structure at 2.8-A resolution of cytochrome c oxidase from				
RT	Paracoccus denitrificans.";				
RL	Nature 376:660-669(1993).				
CC	-1- CATALYTIC ACTIVITY: 4 FERROCYTOCHROME C + O(2) = 2 H(2)O +				
CC	4 FERROCYTOCHROME C.				
CC	-1- SUBCELLULAR LOCATION: INTEGRAL MEMBRANE PROTEIN. INNER MEMBRANE.				
CC	-1- SIMILARITY: BELONGS TO THE CYTOCHROME C OXIDASE SUBUNIT 3 FAMILY.				
CC	This SWISS-PROT entry is copyright. It is produced through a collaboration				
CC	between the Swiss Institute of Bioinformatics and the EMBL Outstation -				
CC	the European Bioinformatics Institute. There are no restrictions on its				
CC	use by non-profit institutions as long as its content is in no way				
CC	modified and this statement is not removed. Usage by and for commercial				
CC	entities requires a license agreement. (See <a href="http://www.isb-sib.ch/announce/">http://www.isb-sib.ch/announce/</a>				
CC	or send an email to <a href="mailto:license@isb-sib.ch">license@isb-sib.ch</a> ).				
CC	-----				
DR	EMBL; X04406; CAA27995.1; -				
DR	EMBL; X05828; CAA29272.1; -				

34	66	46.5	261	1 COX3_EOUAS	P92481 equus asinu
35	66	46.5	261	1 COX3_FELCA	P48892 felis silve
36	66	46.5	261	1 COX3_HORSE	P48661 equus cabal
37	66	46.5	261	1 COX3_HUMAN	P00414 homo sapien
38	66	46.5	261	1 COX3_PAPHA	O92xx8 papio hamad
39	66	46.5	261	1 COX3_RHIUN	O96065 rhinoceros
40	65	45.8	261	1 COX3_CARAU	O96133 carassius a
41	65	45.8	261	1 COX3_CYPCA	P15952 cyprinus ca
42	65	45.8	261	1 COX3_HALGR	P38597 halictetus
43	65	45.8	261	1 COX3_HIPAM	O92425 hippopotamu
44	65	45.8	261	1 COX3_LATCH	O03170 latimeria c
45	65	45.8	261	1 COX3_PONPA	P92696 pongo pygma

DR PIR: S03807; S03807.  
 DR PIR: A24371; A24371.  
 DR HSSP: P00415; 10CC.  
 DR InterPro: IPR000298; Cytochrome\_c\_III.  
 DR Pfam: PF00510; COX3; 1.  
 DR PROSITE: PS50253; COX3; 1.  
 KW Oxidoreductase; Transmembrane; Inner membrane.  
 FT INIT\_MET 0 0  
 FT DOMAIN 1 14 CYTOPLASMIC.  
 FT TRANSEM 15 35  
 FT DOMAIN 36 47 PERIPLASMIC.  
 FT TRANSEM 48 76  
 FT DOMAIN 77 78 CYTOPLASMIC.  
 FT TRANSEM 79 114  
 FT DOMAIN 115 138 PERIPLASMIC.  
 FT TRANSEM 139 165  
 FT DOMAIN 166 167 CYTOPLASMIC.  
 FT TRANSEM 168 196  
 FT DOMAIN 197 202 PERIPLASMIC.  
 FT TRANSEM 203 236  
 FT DOMAIN 237 243 CYTOPLASMIC.  
 FT TRANSEM 244 273  
 SQ SEQUENCE 273 AA; 30655 MW; 0831BD966AE3C7D6 CRC64;

Query Match 73.9%; Score 105; DB 1; Length 273;  
 Best Local Similarity 71.4%; Pred. No. 5e-09;  
 Matches 20; Conservative 4; Mismatches 4; Indels 0; Gaps 0;  
 OY 2 SIFLLVCLIRLRGAMSAKOHVFEEMAA 29  
 Db 224 TIFLVCLIRLRGAMSAKOHVFEEMAA 251

RESULT 2  
 COX3\_AEGCO STANDARD; PRT; 265 AA.  
 AC Q36952;  
 DT 15-JUL-1998 (Rel. 36, Created)  
 DT 15-JUL-1998 (Rel. 36, Last sequence update)  
 DT 15-JUL-1999 (Rel. 38, Last annotation update)  
 DE CYTOCHROME C OXIDASE POLYPEPTIDE III (EC 1.9.3.1).  
 GN COX3.  
 OS Aegilops columnaris.  
 OG Mitochondrion.  
 OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;  
 OC Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae; Pooidae;  
 OC Triticeae; Aegilops.  
 OC NCBI\_TaxID=4493;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RA Ikeda T.M., Tsunewaki K.;  
 RL Submitted (MAR-1996) to the EMBL/GenBank/DBJ databases.  
 CC -1- FUNCTION: SUBUNIT I, II, AND III FORM THE FUNCTIONAL CORE OF  
 CC THE ENZYME COMPLEX.  
 CC -1- CATALYTIC ACTIVITY: 4 FERROCYTOCHROME C + O(2) = 2 H(2)O +  
 CC 4 FERRICYTOCHROME C.  
 CC -1- SIMILARITY: BELONGS TO THE CYTOCHROME C OXIDASE SUBUNIT 3 FAMILY.  
 CC -----  
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -  
 CC the European Bioinformatics Institute. There are no restrictions on its  
 CC use by non-profit institutions as long as its content is in no way  
 CC modified and this statement is not removed. Usage by and for commercial  
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
 CC or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).  
 CC -----  
 CC EMBL: U46766; AAA91211.1; -  
 CC EMBL: U46765; AAA91210.1; -  
 CC InterPro: IPR000298; Cytochrome\_c\_III.  
 CC Pfam: PF00510; COX3; 1.  
 CC PROSITE: PS50253; COX3; 1.  
 KW Oxidoreductase; Mitochondrion; Transmembrane.

SQ SEQUENCE 265 AA; 29326 MW; FF41BF47382FB832 CRC64;  
 Query Match 54.9%; Score 78; DB 1; Length 265;  
 Best Local Similarity 57.1%; Pred. No. 7.4e-05;  
 Matches 16; Conservative 4; Mismatches 8; Indels 0; Gaps 0;  
 OY 2 SIFLLVCLIRLRGAMSAKOHVFEEMAA 29  
 Db 215 TIFLVCLIRLRGAMSAKOHVFEEMAA 242  
 RESULT 3  
 COX3\_WHEAT STANDARD; PRT; 265 AA.  
 AC P15953;  
 DT 01-APR-1990 (Rel. 14, Created)  
 DT 01-AUG-1992 (Rel. 23, Last sequence update)  
 DT 15-JUL-1999 (Rel. 38, Last annotation update)  
 DE CYTOCHROME C OXIDASE POLYPEPTIDE III (EC 1.9.3.1).  
 GN COX3.  
 OS Triticum aestivum (wheat).  
 OG Mitochondrion.  
 OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;  
 OC Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae; Pooidae;  
 OC Triticeae; Triticum.  
 OC NCBI\_TaxID=4565;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RC STRAIN=CV. CAPITOLE;  
 RX MEDLINE=90182728; PubMed=2155710;  
 RA Gualberto J.M., Domon C., Weil J.H., Grienenberger J.M.;  
 RT "Structure and transcription of the gene coding for subunit 3 of  
 RT cytochrome oxidase in wheat mitochondria."  
 RL Curr. Genet. 17:41-47(1990).  
 RN [2]  
 RP SEQUENCE FROM N.A.  
 RC STRAIN=CV. CAPITOLE; TISSUE=Isolated plantlets;  
 RX MEDLINE=92269300; PubMed=1588598;  
 RA Karlovsky P., Fartmann B.;  
 RT "Genetic code and phylogenetic origin of comycetous mitochondria."  
 RL J. Mol. Evol. 34:254-256(1992).  
 RN [3]  
 RP RNA EDITING.  
 RA MEDLINE=90326496; PubMed=1695731;  
 RX Gualberto J.M., Weil J.H., Grienenberger J.M.;  
 RT "Editing of the wheat coxIII transcript: evidence for twelve C to U  
 RT and one U to C conversions and for sequence similarities around  
 RT editing sites."  
 RL Nucleic Acids Res. 18:3771-3776(1990).  
 CC -1- FUNCTION: SUBUNIT I, II, AND III FORM THE FUNCTIONAL CORE OF  
 CC THE ENZYME COMPLEX.  
 CC -1- CATALYTIC ACTIVITY: 4 FERROCYTOCHROME C + O(2) = 2 H(2)O +  
 CC 4 FERRICYTOCHROME C.  
 CC -1- SIMILARITY: BELONGS TO THE CYTOCHROME C OXIDASE SUBUNIT 3 FAMILY.  
 CC -1- CAUTION: POSITIONS 82, 86, 97, 104, 105, 138, 141, 171, 176, 189,  
 CC 252 AND 255 ARE MODIFIED BY RNA EDITING.  
 CC -----  
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -  
 CC the European Bioinformatics Institute. There are no restrictions on its  
 CC use by non-profit institutions as long as its content is in no way  
 CC modified and this statement is not removed. Usage by and for commercial  
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
 CC or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).  
 CC -----  
 CC EMBL: X52539; GAA36775.1; -  
 CC EMBL: X15944; GAA34071.1; ALT\_SEQ.  
 CC PIR: S10331; OTW3M.  
 CC HSSP: P00415; 10CC.  
 CC Mendel: 10988; Triae; cox3; 1.  
 CC InterPro: IPR000298; Cytochrome\_c\_III.  
 CC Pfam: PF00510; COX3; 1.

RC STRAIN=WISTAR;

RX MEDLINE=65022618; PubMed=6091655;  
 RA Pepe G., Holtrop M., Gadaleta G., Kroon A.M., Cantatore P.,  
 RA Gallenzi R., de Benedetto C., Quagliariello C., Sbisa E.,  
 RA Saccocc C.;  
 RT "Non-random patterns of nucleotide substitutions and codon strategy  
 RT in the mammalian mitochondrial genes coding for identified and  
 RT unidentified reading frames";  
 RL Biochem. Int. 6:553-563(1983).  
 RP [3]  
 RP SEQUENCE FROM N.A.  
 RA STRAIN=SPRAGUE-DAWLEY; TISSUE=Liver;  
 RA Grosskopf R., Feldmann H.;  
 RT "Analysis of a DNA segment from rat liver mitochondria containing the  
 RT genes for the cytochrome oxidase subunits I, II, III, ATPase subunit  
 RT 6, and several tRNA genes";  
 RL Curr. Genet. 4:151-158(1981).  
 RN [4]  
 RP SEQUENCE OF 1-10.  
 RC STRAIN=WISTAR; TISSUE=Liver;  
 RX MEDLINE=95324529; PubMed=7601105;  
 RA Schaeffer H., Noack H., Hahangk W., Brandt U., von Jagow G.;  
 RT "Cytochrome c oxidase in developing rat heart. Enzymic properties and  
 RT amino-terminal sequences suggest identity of the fetal heart and the  
 RT adult liver isoform";  
 RL Eur. J. Biochem. 230:235-241(1995).  
 CC -1- FUNCTION: SUBUNIT I, II, AND III FORM THE FUNCTIONAL CORE OF  
 CC THE ENZYME COMPLEX.  
 CC -1- CATALYTIC ACTIVITY: 4 FERROCYTOCHROME C + O(2) = 2 H(2)O +  
 CC 4 FERRICYTOCHROME C.  
 CC -1- SIMILARITY: BELONGS TO THE CYTOCHROME C OXIDASE SUBUNIT 3 FAMILY.  
 CC -----  
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -  
 CC the European Bioinformatics Institute. There are no restrictions on its  
 CC use by non-profit institutions as long as its content is in no way  
 CC modified and this statement is not removed. Usage by and for commercial  
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
 CC or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).  
 CC -----  
 DR EMBL: J01435; AAD15020.1; -;  
 DR EMBL: M27315; AAB00994.1; -;  
 DR EMBL: X14848; CAA32960.1; -;  
 DR PIR: S04753; S04753.  
 DR HSSP: P00415; 10CC.  
 DR InterPro: IPR000298; Cytc\_oxdse\_III.  
 DR Pfam: PF00510; COX3; 1.  
 DR PROSITE: PS50253; COX3; 1.  
 KW Oxidoreductase; Mitochondrion; Transmembrane.  
 KM Oxidoreductase; Mitochondrion; Transmembrane.  
 FT INVT MET 1 0  
 FT CONFLICT 1 0  
 SQ SEQUENCE 260 AA; 29739 MW; AAD4CB2B777C3D3 CRC64;

Query Match 52.8%; Score 75; DB 1; Length 260;  
 Best Local Similarity 53.6%; Pred. No. 0.00021;  
 Matches 15; Conservative 5; Mismatches 8; Indels 0; Gaps 0;

QY 2 SIFLVCLIRIRGAMSAKOHVFEMAA 29  
 DB 211 SIFLVCLIRIRGAMSAKOHVFEMAA 238  
 I |||:||||: ||| ||| ||| |||

RESULT 7  
 COX3\_MOUSE STANDARD; PRT; 261 AA.  
 AC P00416;  
 DT 21-JUL-1986 (Rel. 01, Created)  
 DT 21-JUL-1986 (Rel. 01, Last sequence update)  
 DT 20-AUG-2001 (Rel. 40, Last annotation update)  
 DE CYTOCHROME C OXIDASE POLYPEPTIDE III (EC 1.9.3.1).  
 GN MTCO3 OR MT-CO3 OR COIII.  
 OS Mus musculus (Mouse).  
 OG Mitochondrion.

CC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 CC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.  
 OX NCBI\_TaxID=10090;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RX MEDLINE=82137051; PubMed=7332926;  
 RA Bibb M.O., van Elten R.A., Wright C.T., Walberg M.W., Clayton D.A.;  
 RT "Sequence and gene organization of mouse mitochondrial DNA";  
 RL Cell 26:167-180(1981).  
 CC -1- FUNCTION: SUBUNIT I, II, AND III FORM THE FUNCTIONAL CORE OF  
 CC THE ENZYME COMPLEX.  
 CC -1- CATALYTIC ACTIVITY: 4 FERROCYTOCHROME C + O(2) = 2 H(2)O +  
 CC 4 FERRICYTOCHROME C.  
 CC -1- SIMILARITY: BELONGS TO THE CYTOCHROME C OXIDASE SUBUNIT 3 FAMILY.  
 CC -----  
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -  
 CC the European Bioinformatics Institute. There are no restrictions on its  
 CC use by non-profit institutions as long as its content is in no way  
 CC modified and this statement is not removed. Usage by and for commercial  
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
 CC or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).  
 CC -----  
 DR EMBL: J01420; AAB48650.1; -;  
 DR EMBL: V00711; CAA24090.1; ALT\_TERM.  
 DR PIR: A00484; OTMS3.  
 DR HSSP: P00415; 10CC.  
 DR (MGD: MGI:102502; mt-co3.  
 DR InterPro: IPR000298; mt-co3.  
 DR Pfam: PF00510; COX3; 1.  
 DR PROSITE: PS50253; COX3; 1.  
 KW Oxidoreductase; Mitochondrion; Transmembrane.  
 SQ SEQUENCE 261 AA; 29937 MW; DAF2EA445983D34E CRC64;

Query Match 52.8%; Score 75; DB 1; Length 261;  
 Best Local Similarity 53.6%; Pred. No. 0.00021;  
 Matches 15; Conservative 5; Mismatches 8; Indels 0; Gaps 0;

QY 2 SIFLVCLIRIRGAMSAKOHVFEMAA 29  
 DB 212 SIFLVCLIRIRGAMSAKOHVFEMAA 239  
 I |||:||||: ||| ||| ||| |||

RESULT 8  
 COX3\_PETMA STANDARD; PRT; 261 AA.  
 AC O35539;  
 DT 15-JUL-1998 (Rel. 36, Created)  
 DT 15-JUL-1998 (Rel. 36, Last sequence update)  
 DT 15-JUL-1999 (Rel. 38, Last annotation update)  
 DE CYTOCHROME C OXIDASE POLYPEPTIDE III (EC 1.9.3.1).  
 GN COXIII OR COIII.  
 OS Petromyzon marinus (Sea lamprey).  
 OG Mitochondrion.  
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Hyperoartia;  
 OC Petromyzontiformes; Petromyzontidae; Petromyzon.  
 OX NCBI\_TaxID=7757;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RX MEDLINE=95229067; PubMed=7713438;  
 RA Lee W.J., Koehler T.D.;  
 RT "Complete sequence of a sea lamprey (Petromyzon marinus)  
 RT mitochondrial genome: early establishment of the vertebrate genome  
 RT organization";  
 RL Genetics 139:873-887(1995).  
 CC -1- FUNCTION: SUBUNIT I, II, AND III FORM THE FUNCTIONAL CORE OF  
 CC THE ENZYME COMPLEX.  
 CC -1- CATALYTIC ACTIVITY: 4 FERROCYTOCHROME C + O(2) = 2 H(2)O +  
 CC 4 FERRICYTOCHROME C.  
 CC -1- SIMILARITY: BELONGS TO THE CYTOCHROME C OXIDASE SUBUNIT 3 FAMILY.  
 CC -----  
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration



CC between the Swiss Institute of Bioinformatics and the EMBL outstation  
CC the European Bioinformatics Institute. There are no restrictions on lists  
CC use by non-profit institutions as long as its content is in no way  
CC modified and this statement is not removed. Usage by and for commercial  
CC entities requires a license agreement (See [http://www.isb-sib.ch/announcement/  
CC or send an email to license@isb-sib.ch](http://www.isb-sib.ch/announcement/or-send-an-email-to-license@isb-sib.ch)).

DR EMBL U11880: AAB08744.1; -.  
DR InterPro: IPR000298; Cyto\_c-oxdase-III.  
DR Pfam: PF00510; COX3; 1.  
DR PROSITE: PS50253; COX3; 1.  
KW Oxidoreductase; Mitochondrion; Transmembrane.  
SO SEQUENCE 261 AA; 2917 MW; 314B955DDC9A7B7E CRC64;

Query Match	52.8%	Score	75	DB	1	Length	261
Best Local Similarity	53.6%	Pred.	No.	0.00021			
Matches	15	Conservative	5	Mismatches	8	Indels	0
						Gaps	0

QY 2 SIFLLVCLIRILRGAMSAKOHVFEEMAA 29  
|:|:| |:|:| |:|:| |:|:| |  
Db 212 SLFLLTCLLRHLQYHFTSKHNFGEFAAA 239

RESULT	9	
COX3_MAIZE		
ID	COX3_MAIZE	STANDARD;
		PRT;
		265 AA

DT 01-NOV-1995 (Rel. 32, Last sequence update)  
DT 15-JUL-1999 (Rel. 38, Last annotation update)  
DE CYTOCHROME C OXIDASE POLYPEPTIDE III (EC 1.9.3.1)  
NOV2

OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;  
OC Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae; PACC clade;  
OC Panicoidae; Andropogoneae; Zea.

RP SEQUENCE FROM N.A.  
RX MEDLINE=69041586; PubMed=2847119;  
RA McCarthy D.M., Helman G.L., Hauswirth W.W.;  
RT "Nucleotide sequence of the zea mays mitochondrial  
RT subunit III gene".  
RL Nucleic Acids Res. 16:9873-9873(1988).  
not

RP SEQUENCE FROM N.A.  
RC SPRAIN-CV. B37N.  
RA Kaleikau E. K., Walbot V.:  
Submitted (APR-1992) to the EMBL/Genbank/DBJ databases.  
CC -1- FUNCTION: SUBUNIT I, II, AND III FORM THE FUNCTIONAL CORE OF  
THE ENZYME COMPLEX.  
CC -1- CATALYTIC ACTIVITY: 4 FERROCYTOCHROME C + O(2) = 2 H(2)O +  
4 FERROCYTOCHROME C.  
CC -1- SIMILARITY: BELONGS TO THE CYTOCHROME C OXIDASE SUBUNIT 3 FAMILY.

CC This SWISSPORT entry is copyright. It is produced through a collaboration  
CC between the Swiss Institute of Bioinformatics and the EMBL Outstation  
CC at the European Bioinformatics Institute. There are no restrictions on lists  
CC use by non-profit institutions as long as its content is in no way  
CC modified and this statement is not removed. Usage by and for commercial  
CC entities requires a license agreement (see [http://www.isb-sib.ch/announcement/  
CC or\\_send\\_an\\_email\\_to\\_license@isb-sib.ch](http://www.isb-sib.ch/announcement/or_send_an_email_to_license@isb-sib.ch)).

DR EMBL: X12728; CA31221.1; -.  
DR EMBL: K53055; CA37222.1; -.  
DR HSSP: P00415; 100C.  
DR Mz1zDB: 69221; -.  
DR Mendel: 5492; ZEMMA:cox3.1.  
DR InterPro: IPR00298; CytC\_oxdase\_III  
DR Pfam: P00510; COX3; 1.  
DR PROSITE: PS50253; COX3; 1.

KW Oxidoreductase; Mitochondrion; Transmembrane.  
SQ SEQUENCE 265 AA; 29328 MW; B10D10921745180C CRC64;

Query Match	52.1%	Score 74;	DB 1;	Length 265;
Best Local Similarity	53.6%	Pred. No. 0.00031;		
Matches 15; Conservative	5;	Mismatches 8;	Indels 0;	Gaps 0;

QY    2   SIFLLVCLIRLRGAMSAKQHVFEMAA    29  
         ::||| ||    :: | |||| |  
Db    215 TLFLLVCGIRQYLGHILTKKHNVGFEEAA    242

RESULT	10	
COX3_ORYSA		
ID	COX3_ORYSA	STANDARD;
PLAGE		PRT;
		265 AA

DT 01-DEC-1992 (Rel. 24, Last sequence update)  
DT 15-JUL-1999 (Rel. 38, Last annotation update)  
DE CYTOCHROME C OXIDASE POLYPEPTIDE I II (EC 1.9.3.1),  
COX3

OC Eukaryota: Viridiplantae: Streptophyta; Embryophyta; Tracheophyta  
OC Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae;  
OC Ehrhartoideae; Oryzeae; Oryza.

RP SEQUENCE FROM N.A.  
RC STRAIN=CV. INDICA-IR36;  
RX MEDLINE=90221829; PubMed=2158075;

RT "Sequence of the rice mitochondrial gene for cytochrome oxidase subunit 3.";  
RT  
RL Nucleic Acids Res. 18:371-371(1990).  
188

RA Liu A.W., Narayanan K.K., Andre C.P., Kaleikau E.K., Walbot V.;  
RT "Co-transcription of orf25 and coxIII in rice mitochondria."; *J.*  
RL *Curr. Genet.* 21:507-513 (1992).

CC -1- CATALYTIC ACTIVITY: 4 FERROCYTOCHROME C + O(2) = 2 H(2)O +  
4 FERRICCYTOCHROME C.  
CC -1- SIMILARITY: BELONGS TO THE CYTOCHROME C OXIDASE SUBUNIT 3 FAMILY

CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
CC between the Swiss Institute of Bioinformatics and the EMBL outstation  
CC at the European Bioinformatics Institute. There are no restrictions on its  
CC use by non-profit institutions as long as its content is in no way  
CC modified and this statement is not removed. Usage by and for commercial  
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
CC or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).

DR EMBL: M17040; CAA34898.1; -.  
DR EMBL: M74241; AAA66048.1; -.  
DR PIR: J00166; ATR23M.  
DR PIR: S20801; S20801.  
DR PIR: S29748; S29748.  
DR HSSP: P00415; LOCC.  
DR Mendel: 5490; ORSs:cox3.1.  
DR InterPro: IPR000298; CytC\_oxase\_III.  
DR Pfam: PF00510; COX3.1.  
DR PROSITE: PS50253; COX3.1.  
DR Oxidoreductase; Mitochondrion; Transmembrane.  
KW EC:FB453E097C1D CRC64; EC:FB453E097C1D CRC64;  
SEQUENCE 265 aa; 29358 MW; EC:FB453E097C1D CRC64;

Query Match	52.18; Score 74; DB 1; Length 265;
Best Local Similarity	53.68; Pred. No. 0.00031;



RP SEQUENCE OF 179-262 FROM N.A.  
 RC STRAIN-BRETAGNE;  
 RX MEDLINE-88212147; PubMed-3130291;  
 RA Garesse R.;  
 RT "Drosophila melanogaster mitochondrial DNA: gene organization and  
 evolutionary considerations.";  
 RL Genetics 118:649-663(1988).  
 CC  
 CC -1- FUNCTION: SUBUNIT I, II, AND III FORM THE FUNCTIONAL CORE OF  
 CC THE ENZYME COMPLEX.  
 CC -1- CATALYTIC ACTIVITY: 4 FERROCYTOCHROME C + O(2) = 2 H(2)O +  
 CC 4 FERRICYTOCHROME C.  
 CC  
 CC -1- SIMILARITY: BELONGS TO THE CYTOCHROME C OXIDASE SUBUNIT 3 FAMILY.  
 CC  
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -  
 CC the European Bioinformatics Institute. There are no restrictions on its  
 CC use by non-profit institutions as long as its content is in no way  
 CC modified and this statement is not removed. Usage by and for commercial  
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
 CC or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).  
 CC  
 CC EMBL: J01404; AAB59243.1; -  
 CC EMBL: M37275; AAB69708.1; -  
 CC EMBL: U37541; AAC47816.1; -  
 CC PIR: B93477; OTFE3Y.  
 CC PIR: S02250; S02250.  
 CC HSSP: P00415; 10CC.  
 CC FlyBase: FBgn0013676; mt:CoIII.  
 CC InterPro: IPR000298; Cytc\_oxdse\_III.  
 CC Pfam: PF00510; COX3; 1.  
 CC PROSITE: PS50253; COX3; 1.  
 CC Oxidoreductase; Mitochondrion; Transmembrane.  
 KW  
 SQ SEQUENCE 262 AA; 30064 MW; 68D5AEFC2297C130 CRC64;

Query Match 50.0%; Score 71; DB 1; Length 262;  
 Best Local Similarity 57.7%; Pred. No. 0.00089;  
 Matches 15; Conservative 1; Mismatches 10; Indels 0; Gaps 0;

OY 4 FLVCLIRLRGMSAKOHVGFEMAA 29  
 DB 215 FLVCLIRLRHNFHFSKNHGFEMAA 240

RESULT 14  
 COX3\_DROYA STANDARD; PRT; 262 AA.  
 AC P00418;  
 DT 21-JUL-1986 (Rel. 01, Created)  
 DT 15-JUL-1986 (Rel. 01, Last sequence update)  
 DT 15-JUL-1999 (Rel. 38, Last annotation update)  
 DE CYTOCHROME C OXIDASE POLYPEPTIDE III (EC 1.9.3.1).  
 GN COIII.  
 OS Drosophila yakuba (Fruit fly).  
 OG Mitochondrion.  
 OC Eukaryota; Metazoa; Arthropoda; Tracheata; Hexapoda; Insecta;  
 OC Pterygota; Neoptera; Endopterygota; Diptera; Brachycera; Muscomorpha;  
 OC Ephydroidea; Drosophilidae; Drosophila.  
 OX NCBI\_Taxid=7245;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RC STRAIN-2317.6 IVORY COAST;  
 RX MEDLINE-86089137; PubMed-3001325;  
 RA Clary D.O.; Wolstenholme D.R.;  
 RT "The mitochondrial DNA molecular of Drosophila yakuba: nucleotide  
 sequence, gene organization, and genetic code.";  
 RL J. Mol. Evol. 22:252-271(1985).  
 RN [2]  
 RP SEQUENCE FROM N.A.  
 RX MEDLINE-83246544; PubMed-6306579;  
 RA Clary D.O.; Wolstenholme D.R.;  
 RT "Nucleotide sequence of a segment of Drosophila mitochondrial DNA  
 that contains the genes for cytochrome c oxidase subunits II and III

RT and ATPase subunit 6.";  
 RL Nucleic Acids Res. 11:4211-4227(1983).  
 CC  
 CC -1- FUNCTION: SUBUNIT I, II, AND III FORM THE FUNCTIONAL CORE OF  
 CC THE ENZYME COMPLEX.  
 CC -1- CATALYTIC ACTIVITY: 4 FERROCYTOCHROME C + O(2) = 2 H(2)O +  
 CC 4 FERRICYTOCHROME C.  
 CC  
 CC -1- SIMILARITY: BELONGS TO THE CYTOCHROME C OXIDASE SUBUNIT 3 FAMILY.  
 CC  
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -  
 CC the European Bioinformatics Institute. There are no restrictions on its  
 CC use by non-profit institutions as long as its content is in no way  
 CC modified and this statement is not removed. Usage by and for commercial  
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
 CC or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).  
 CC  
 CC EMBL: X03240; CAA26990.1; -  
 CC EMBL: X00924; CAA25443.1; -  
 CC PIR: A00465; OTFE3.  
 CC PIR: P25797; P25797.  
 CC HSSP: P00415; 10CC.  
 CC FlyBase: FBgn0013181; Dyak\mt:CoIII.  
 CC InterPro: IPR000298; Cytc\_oxdse\_III.  
 CC Pfam: PF00510; COX3; 1.  
 CC PROSITE: PS50253; COX3; 1.  
 CC Oxidoreductase; Mitochondrion; Transmembrane.  
 KW  
 SQ SEQUENCE 262 AA; 30039 MW; EDDC07F573666730 CRC64;

Query Match 50.0%; Score 71; DB 1; Length 262;  
 Best Local Similarity 57.7%; Pred. No. 0.00089;  
 Matches 15; Conservative 1; Mismatches 10; Indels 0; Gaps 0;

OY 4 FLVCLIRLRGMSAKOHVGFEMAA 29  
 DB 215 FLVCLIRLRHNFHFSKNHGFEMAA 240

RESULT 15  
 COX3\_GADMO STANDARD; PRT; 261 AA.  
 AC P55777;  
 DT 01-NOV-1997 (Rel. 35, Created)  
 DT 01-NOV-1997 (Rel. 35, Last sequence update)  
 DT 20-AUG-2001 (Rel. 40, Last annotation update)  
 DE CYTOCHROME C OXIDASE POLYPEPTIDE III (EC 1.9.3.1).  
 GN COXIII OR COIII.  
 OS Gadus morhua (Atlantic cod).  
 OG Mitochondrion.  
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 OC Actinopterygii; Neopterygii; Teleostei; Euteleostei; Neoteleostei;  
 OC Acanthomorpha; Paracanthopterygii; Gadiformes; Gadidae;  
 OC Gadus.  
 OX NCBI\_Taxid=8049;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RC STRAIN-NORWEGIAN COASTAL 1;  
 RX MEDLINE-96414925; PubMed-8817926;  
 RA Johansen S.; Bakke I.;  
 RT "The complete mitochondrial DNA sequence of Atlantic cod (Gadus  
 morhua): relevance to taxonomic studies among codfishes.";  
 RL Mol. Mar. Biol. Biotechnol. 5:203-214(1996).  
 CC  
 CC -1- FUNCTION: SUBUNIT I, II, AND III FORM THE FUNCTIONAL CORE OF  
 CC THE ENZYME COMPLEX.  
 CC -1- CATALYTIC ACTIVITY: 4 FERROCYTOCHROME C + O(2) = 2 H(2)O +  
 CC 4 FERRICYTOCHROME C.  
 CC  
 CC -1- SIMILARITY: BELONGS TO THE CYTOCHROME C OXIDASE SUBUNIT 3 FAMILY.  
 CC  
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -  
 CC the European Bioinformatics Institute. There are no restrictions on its  
 CC use by non-profit institutions as long as its content is in no way  
 CC modified and this statement is not removed. Usage by and for commercial



GenCore version 4.5  
Copyright (c) 1993 - 2000 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: January 10, 2002, 02:04:05 ; Search time 32.55 Seconds  
(without alignments)  
130.319 Million cell updates/sec

Title: US-09-712-768-8  
Perfect score: 142  
Sequence: 1 DSIFLVCLIRITLGAMSAKOHGFEMAA 29

Scoring table: BLOSUM62  
Gapop 10.0 , Gapext 0.5

Searched: 473505 seqs, 146272329 residues  
Total number of hits satisfying chosen parameters: 473505

Minimum DB seq length: 0  
Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%  
Maximum Match 100%  
Listing first 45 summaries

Database :  
1: SPREMBL\_17.\*  
2: sp-archaea.\*  
3: sp-bacteria.\*  
4: sp-fungi.\*  
5: sp-human.\*  
6: sp-invertebrate.\*  
7: sp-mammal.\*  
8: sp-mhc.\*  
9: sp-organellar.\*  
10: sp-phage.\*  
11: sp-plant.\*  
12: sp-rodent.\*  
13: sp-virus.\*  
14: sp-vertebrate.\*  
15: sp-unclassified.\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

## SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	87	61.3	290	2	Q9A304
2	86	60.6	298	2	Q9RM95
3	79	55.6	274	8	Q9G874
4	79	55.6	281	8	Q9G986
5	78	54.9	261	8	Q9Z231
6	77	54.2	272	8	Q9Z208
7	76	53.5	219	8	Q9S948
8	75	52.8	227	8	Q9B332
9	75	52.8	230	8	Q9B330
10	75	52.8	232	8	Q9B327
11	75	52.8	232	8	Q9B321
12	75	52.8	233	8	Q9G3G9
13	75	52.8	233	8	Q9G3G8
14	75	52.8	233	8	Q9G3G7
15	75	52.8	233	8	Q9G3G5
16	75	52.8	233	8	Q9G3F9
17	75	52.8	233	8	Q9G3F4
18	75	52.8	233	8	Q9G3F0
19	75	52.8	233	8	Q9G3E6

20	75	52.8	233	8	Q9G3E5	Q9G3E5	peromyscus
21	75	52.8	233	8	Q9G3D0	Q9G3D0	peromyscus
22	75	52.8	233	8	Q9G3D9	Q9G3D9	peromyscus
23	75	52.8	233	8	Q9G3D6	Q9G3D6	peromyscus
24	75	52.8	233	8	Q9G3D5	Q9G3D5	peromyscus
25	75	52.8	233	8	Q9G2H4	Q9G2H4	peromyscus
26	75	52.8	233	8	Q9G150	Q9G150	peromyscus
27	75	52.8	233	8	Q9G147	Q9G147	peromyscus
28	75	52.8	233	8	Q9G0V6	Q9G0V6	peromyscus
29	75	52.8	233	8	Q9G0V5	Q9G0V5	peromyscus
30	75	52.8	233	8	Q9G0V4	Q9G0V4	peromyscus
31	75	52.8	233	8	Q9B329	Q9B329	peromyscus
32	75	52.8	233	8	Q9B323	Q9B323	peromyscus
33	75	52.8	233	8	Q9B319	Q9B319	peromyscus
34	75	52.8	233	8	Q9B318	Q9B318	peromyscus
35	75	52.8	233	8	Q9B0S6	Q9B0S6	peromyscus
36	75	52.8	244	8	Q9M1V9	Q9M1V9	mesocricetu
37	75	52.8	261	8	Q9M0M4	Q9M0M4	mus musculu
38	75	52.8	261	8	Q9MDE6	Q9MDE6	mus musculu
39	75	52.8	262	8	Q9TBV9	Q9TBV9	mus musculu
40	75	52.8	264	8	Q9T9H6	Q9T9H6	halocythia
41	74	52.1	260	8	Q9MNMJ0	Q9MNMJ0	halocythia
42	74	52.1	261	8	Q9TA05	Q9TA05	lampectra fl
43	73	51.4	259	8	Q9MNMJ9	Q9MNMJ9	galathea
44	73	51.4	260	8	Q36906	Q36906	arabacta lix
45	73	51.4	261	8	Q9X131	Q9X131	vidua chaly

## ALIGNMENTS

RESULT 1  
ID Q9A304 PRELIMINARY; PRT: 290 AA.  
AC Q9A304;  
DT 01-JUN-2001 (TREMBLrel. 17, Created)  
DT 01-JUN-2001 (TREMBLrel. 17, Last sequence update)  
DT 01-JUN-2001 (TREMBLrel. 17, Last annotation update)  
DE CYTOCHROME C OXIDASE, SUBUNIT III.  
GN CC3402.  
OS Caulobacter crescentus.  
OC Bacteria; Proteobacteria; alpha subdivision; Caulobacter group;  
OC Caulobacter.  
OX NCBI\_Taxid=69394;  
RN [1]  
RP SEQUENCE FROM N.A.  
RX MEDLINE=21173698; PubMed=11259647;  
RA Nierman W.C., Feldblyum T.V., Laub M.T., Paulsen I.T., Nelson K.E.,  
RA Eisen J., Heidelberg J.F., Alley M.R.K., Ohta N., Maddock J.R.,  
RA Potocka I., Nelson W.C., Newton A., Stephens C., Phadke N.D., Ely B.,  
RA DeBoy R.T., Dodson R.J., Durkin A.S., Gwinn M.L., Haft D.H.,  
RA Kolonay J.F., Smit J., Craven M.B., Khouri H., Shetty J., Berry K.,  
RA Salzberg S.L., Tran K., Wolf A., Vamathevan J., Ermolaeva M., White O.,  
RA Utterback T., Venter J.C., Shapiro L., Fraser C.M.,  
RT "Complete genome sequence of Caulobacter crescentus."  
RL Proc. Natl. Acad. Sci. U.S.A. 98:4136-4141(2001).  
DR EMBL: AE006000; AAK25364.1; .  
DR TIGR: CC3402; .  
KW Complete proteome.  
SQ SEQUENCE 290 AA; 32229 MW; 47016BD7090AB972 CRC64;

Query Match 61.3%; Score 87; DB 2; Length 290;  
Best Local Similarity 53.6%; Pred. No. 3.5e-06;  
Matches 15; Conservative 8; Mismatches 5; Indels 0; Gaps 0;

QY 2 SIFLVCLIRITLGAMSAKOHGFEMAA 29  
Db 237 TFLIVCLIRITMGALTPQKHGFEMAA 264

RESULT 2  
Q9RM95

AD	OC9RM95	PRELIMINARY;	PRT;	298 AA.
AC	OC9RM95;			
DT	01-MAY-2000 (TREMBLrel. 13, Created)			
DT	01-MAY-2000 (TREMBLrel. 13, Last sequence update)			
DT	01-JUN-2001 (TREMBLrel. 17, Last annotation update)			
DE	CYCLOHOME C OXIDASE SUBUNIT III (EC 1.9.3.1) (COXC).			
GN	COXC.			
OS	Bradyrhizobium japonicum.			
OC	Bacteria; Proteobacteria; alpha subdivision; Rhizobiaceae group;			
OC	Bradyrhizobium group; Bradyrhizobium.			
OX	NCBI_TaxId=375;			
RN	[1]			
RP	SEQUENCE FROM N.A.			
RC	STRAIN=110SPC4;			
RA	Rossmann R., Lofterer H., Rossi P., Hennecke H.;			
RT	"Factors involved in biosynthesis of active cytochrome a3 encoded by			
RL	the COXB4EFG gene cluster from Bradyrhizobium japonicum.";			
RT	Submitted (MAY-1999) to the EMBL/Genbank/DBDj databases.			
RN	[2]			
RP	SEQUENCE FROM N.A.			
RC	STRAIN=USDA110SPC4;			
RX	MEDLINE=96422470; PubMed=8825087;			
RA	Muller P., Ahrens K., Keller T., Klauke A.;			
RT	"A Taphoa insertion within the Bradyrhizobium japonicum sips gene,			
RT	homologous to prokaryotic signal peptidases, results in extensive			
RT	changes in the expression of PBM-specific nodulins of infected soybean			
RL	(glycine max) cells.";			
RL	Mol. Microbiol. 18:831-840(1995).			
RN	[3]			
RP	SEQUENCE FROM N.A.			
RC	STRAIN=USDA110SPC4;			
RA	Mueller P.;			
RT	"Extended sequencing of a DNA fragment of B. japonicum adjacent to the			
RT	cox operon.";			
RL	Submitted (JUN-2000) to the EMBL/Genbank/DBDj databases.			
DR	EMBL; AJ242592; CAB56823.1; -.			
DR	EMBL; U33883; AAF7881.1; -.			
DR	InterPro; IPR000298; CytC_oxdase_III.			
DR	Pfam; PF00510; COX3; 1.			
DR	PROSITE; PS05253; COX3; 1.			
KM	Oxidoreductase.			
SO	SEQUENCE 298 AA; 32985 MW; 6AE3D9E5AC2D6C63 CRC64;			

```

Query Match          60.6%; Score 86; DB 2; Length 298;
Best Local Similarity 60.7%; Pred. No. 5.2e-06;
Matches 17; Conservative 3; Mismatches 8; Indels 0; Gaps 0.

OY      2 SIFLLVCLIRIDRGAMSAKOHVGFEMAA 29
          :||||| | | : |||: || | |
DB      239 TIFLLVCLIFRAYAGHTPKOHGLGFEPAA 266

RESULT  3
ID      09G874      PRELIMINARY;      PRT;      274 AA.
AC      09G874;
DT      01-MAR-2001 (TREMBLrel. 16, Created)
DT      01-MAR-2001 (TREMBLrel. 16, Last sequence update)
DT      01-JUN-2001 (TREMBLrel. 17, Last annotation update)
DE      CYTOCHROME C OXIDASE SUBUNIT 3 (Ec 1.9.3.1).
GN      COX3.
OS      Malawimonas jakobiformis.
OC      Mitochondrion.
OC      Eukaryota; Malawimonadidae; Malawimonas.
OX      NCBI_TaxID=136089;
RN      [1]
RP      SEQUENCE FROM N.A.
RA      Burger G., O'Kelly C.J., Gray M.W., Lang B.F.;
RT      "Comparative analysis of mitochondrial genomes of the ancient jakobid
RT      protists.",
RL      Submitted (AUG-2000) to the EMBL/GenBank/DBJ databases.
DR      EMBL; AF295546; AAG13699.1; -.

```

DR InterPro: IPR000298; Gyc\_oxdse\_III.  
DR Pfam: PF00510; COX3; 1.  
DR PROSITE: PS03253; COX3; 1.  
KW Oxidoreductase; Mitochondrion.  
SQ SEQUENCE 274 AA; 31008 MW; 361ADE354C64CE57 CRC64

Query Match	55.68;	Score 79;	DB 8;	Length 274;
Best Local Similarity	46.48;	Pred. No. 6.8e-05;		
Matches 13;	Conservative 9;	Mismatches 6;	Indels 0;	Gaps 0;

QY 2 SIFLLVCLIRILRGAMSAKQHVGFEMAA 29  
:::|::|::: :| | | | | |  
Db 222 TLFLLVCLVRLIKYHFTSKHHGFEEAAM 249

RESULT	4	
099986		
ID	099986	PRELIMINARY;
1		PRT;
000000		281 AA

DT 01-MAY-1999 (TREMBLrel. 10, Created)  
DT 01-MAY-1999 (TREMBLrel. 10, last sequence update)  
DT 01-JUN-2001 (TREMBLrel. 17, last annotation update)  
DE CYTOCHROME C OXIDASE SUBUNIT 3 (EC 1.9.3.1).  
GN COX3.  
OS Porphyra purpurea.  
OG Mitochondrion.  
OC Eukaryota; Rhodophyta; Bangiophyceae; Bangiales; Bangiaceae; Porphyra.  
OX NEBL\_TaxId=2787;  
[1]  
RN SEQUENCE FROM N.A.  
RP Burger G., Saint-Louis D., Gray M.W., Lang B.F.;  
RA "Complete sequence of the mitochondrial DNA of the red alga, Porphyra  
RT purpurea. Inverted repeats, sequence polymorphisms, and cyanobacterial  
RT introns";  
RL Submitted (DEC-1998) to the EMBL/GenBank/DBJ databases.  
DR EMBL; AF114794; AB003114.1; -;  
DR InterPro; IPR000298; CytC\_oxdase\_III.  
DR Pfam; PF00510; COX3; 1.  
DR PROSITE; PSS0253; COX3; 1.  
KW Oxidoreductase; Mitochondrion.  
SQ SEQUENCE 281 AA; 31794 MW; 14C4F376F04901C6 CRC64;

Query Match	55.6%;	Score 79;	DB 8;	Length 281;
Best Local Similarity	50.0%;	Pred No. 6.9e-05;		
Matches 14; Conservative	7;	Mismatches 7;	Indels 0;	Gaps 0;

```
QY      2 SIFLLVCLIRILRGAMSAKQHVGFEAA 29  
       :|||:||||:|::| ||| ||  
Db     231 TIFLAICLIRLKSHTQQHHGFEEAA 258
```

RESULT	5		
09Z231			
ID	09Z231	PRELIMINARY;	PRT; 261 AA.
AC	09Z231;		
DT	01-MAY-1999 (T-Emblrel. 10, Created)		
DT	01-MAY-1999 (T-Emblrel. 10, Last sequence update)		
DT	01-JUN-2001 (T-Emblrel. 17, Last annotation update)		
DE	CYTOCHROME OXIDASE SUBUNIT III (FRAGMENT).		
GN	COIII.		
OS	Corvus frugilegus.		
OS	Mitochondrion.		
OC	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;		
OC	Archosauria; Aves; Neognathae; Passeriformes; Corvidae; Corvus.		
OX	NCBI_TaxID=75140;		
RN	[1]		
RP	SEQUENCE FROM N.A.		
RA	Harlid A., Arnason U.;		
RT	"Analyses of mitochondrial DNA nest ratite birds within the		
RT	Neognathae supporting a neotenuous origin of ratite morphological		
RT	characters.";		

RL Proc. R. Soc. Lond., B, Biol. Sci. 266:305-309(1999).  
 RN [2]  
 RP SEQUENCE FROM N.A.  
 RA Harild A.;  
 RL Submitted (DEC-1998) to the EMBL/Genbank/DBJ databases.  
 DR EMBL: Y18522; CAA77201.1; -  
 DR HSP: P18402; 1FRT.  
 DR Interpro: IPR000298; Cytc\_oxdse\_III.  
 DR Pfam: PF00510; COX3; 1.  
 DR PROSITE: PS50253; COX3; 1.  
 KM Mitochondrion.  
 FT NON\_TER 261  
 SQ SEQUENCE 261 AA; 29532 MW; 5021795FC7B5FD2 CRC64;

Query Match 54.9%; Score 78; DB 8; Length 261;  
 Best Local Similarity 53.6%; Pred. No. 9.4e-05;  
 Matches 15; Conservative 5; Mismatches 8; Indels 0; Gaps 0;

OY 2 STELLVCLIRILRGAMSAKOHVFEMAA 29  
 DB 212 STELLVCLIRILRGAMSAKOHVFEMAA 239

RESULT 6  
 ID 092Z08 PRELIMINARY; PRT; 272 AA.

AC 092Z08;  
 DT 01-MAY-1999 (Tremblrel. 10, Created)  
 DT 01-MAY-1999 (Tremblrel. 10, Last sequence update)  
 DT 01-JUN-2001 (Tremblrel. 17, Last annotation update)  
 DE CYTOCHROME C OXIDASE POLYPEPTIDE III (EC 1.9.3.1).  
 GN COX3.  
 OS Cyanidioschyzon merolae.  
 OC Mitochondrion.  
 OC Eukaryota; Rhodophyta; Bangiophyceae; Cyanidioschyzon.  
 OX NCBI\_TaxID=45157;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RC STRAIN=10D;  
 RC MEDLINE=99030526; PubMed=9801318;  
 RA Onita N., Sato N., Kuroiwa T.;  
 RT "Structure and organization of the mitochondrial genome of the unicellular red alga Cyanidioschyzon merolae deduced from the complete nucleotide sequence."  
 RT Nucleic Acids Res. 26:5190-5198(1998).  
 RL EMBL: D89861; BAA34657.1; -  
 DR HSP: P18402; 1FRT.  
 DR Interpro: IPR000298; Cytc\_oxdse\_III.  
 DR Pfam: PF00510; COX3; 1.  
 DR PROSITE: PS50253; COX3; 1.  
 KM Oxidoreductase; Mitochondrion.  
 SQ SEQUENCE 272 AA; 31047 MW; ACEA7E8BEF1F4B6C CRC64;

Query Match 54.2%; Score 77; DB 8; Length 272;  
 Best Local Similarity 50.0%; Pred. No. 0.00014;  
 Matches 14; Conservative 7; Mismatches 7; Indels 0; Gaps 0;

OY 2 STELLVCLIRILRGAMSAKOHVFEMAA 29  
 DB 222 STELLVCLIRILRGAMSAKOHVFEMAA 249

RESULT 7  
 ID 095948 PRELIMINARY; PRT; 219 AA.

AC 095948;  
 DT 01-FEB-1997 (Tremblrel. 02, Created)  
 DT 01-FEB-1997 (Tremblrel. 02, Last sequence update)  
 DT 01-JUN-2001 (Tremblrel. 17, Last annotation update)  
 DE CYTOCHROME OXIDASE SUBUNIT III (FRAGMENT).  
 GN COXIII.

OS Sepia elegans.  
 OG Mitochondrion.  
 OC Eukaryota; Metazoa; Mollusca; Cephalopoda; Coleoidea; Sepioida;  
 OC Sepiidae; Sepia.  
 OX NCBI\_TaxID=34527;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RA MEDLINE=97159560; PubMed=9007019;  
 RA Bonnaud L., Boucher Rodoni R., Monnerot M.;  
 RT "Phylogeny of cephalopods inferred from mitochondrial DNA sequences."  
 RL Mol. Phylogenet. Evol. 7:44-54(1997).  
 DR EMBL: Y07843; CAA69171.1; -  
 DR Interpro: IPR000298; Cytc\_oxdse\_III.  
 DR Pfam: PF00510; COX3; 1.  
 DR PROSITE: PS50253; COX3; 1.  
 KM Mitochondrion.  
 FT NON\_TER 1  
 SQ SEQUENCE 219 AA; 24743 MW; 2454BP6ECAD15E92 CRC64;

Query Match 53.5%; Score 76; DB 8; Length 219;  
 Best Local Similarity 57.1%; Pred. No. 0.00017;  
 Matches 16; Conservative 3; Mismatches 9; Indels 0; Gaps 0;

OY 2 STELLVCLIRILRGAMSAKOHVFEMAA 29  
 DB 192 STELLVCLIRILRGAMSAKOHVFEMAA 219

RESULT 8  
 ID 09B332 PRELIMINARY; PRT; 227 AA.

AC 09B332;  
 DT 01-JUN-2001 (Tremblrel. 17, Created)  
 DT 01-JUN-2001 (Tremblrel. 17, Last sequence update)  
 DT 01-JUN-2001 (Tremblrel. 17, Last annotation update)  
 DE CYTOCHROME OXIDASE SUBUNIT III (FRAGMENT).  
 GN COXIII.  
 OS Peromyscus slevini (Slevin's mouse).  
 OC Mitochondrion.  
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Sigmodontinae;  
 OC Peromyscus.  
 OX NCBI\_TaxID=44242;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RC STRAIN=2439\_CATALINA;  
 RA Hainer D.J., Riddle B.R., Alvarez-Castaneda S.T.;  
 RT "Evolutionary relationships of white-footed mice (Peromyscus) on islands in the Sea of Cortez, Mexico."  
 RT Submitted (JAN-2001) to the EMBL/Genbank/DBJ databases.  
 RL EMBL: AF343755; AAK32167.1; -  
 DR EMBL: AF343755; AAK32167.1; -  
 KM Mitochondrion.  
 FT NON\_TER 1  
 FT NON\_TER 227  
 SQ SEQUENCE 227 AA; 25524 MW; E2157F384CD68029 CRC64;

Query Match 52.8%; Score 75; DB 8; Length 227;  
 Best Local Similarity 53.6%; Pred. No. 0.00025;  
 Matches 15; Conservative 5; Mismatches 8; Indels 0; Gaps 0;

OY 2 STELLVCLIRILRGAMSAKOHVFEMAA 29  
 DB 200 STELLVCLIRILRGAMSAKOHVFEMAA 227

RESULT 9  
 ID 09B330 PRELIMINARY; PRT; 230 AA.

AC 09B330;  
 DT 01-JUN-2001 (Tremblrel. 17, Created)

DT 01-JUN-2001 (Tremblrel. 17, Last sequence update)  
 DE CYTOCHROME OXIDASE SUBUNIT III (FRAGMENT).  
 GN COIII.  
 OS Peromyscus interparietalis.  
 OG Mitochondrion.  
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Sigmodontinae;  
 OC Peromyscus.  
 NCBI\_TaxID=152342;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RC STRAIN=2438\_SALSIPUEDES;  
 RA Hafner D.J., Riddle B.R., Alvarez-Castaneda S.T.;  
 RT "Evolutionary relationships of white-footed mice (Peromyscus) on  
 islands in the Sea of Cortez, Mexico."  
 RL Submitted (JAN-2001) to the EMBL/GenBank/DBJ databases.  
 DR EMBL: AF343757; AAK32169.1; -  
 KW Mitochondrion.  
 FT NON\_TER 1 1  
 FT 230 230  
 SQ SEQUENCE 230 AA; 25894 MW; DPAACDCEFE6571A CRC64;

Query Match 52.8%; Score 75; DB 8; Length 230;  
 Best Local Similarity 53.6%; Pred. No. 0.00026;  
 Matches 15; Conservative 5; Mismatches 8; Indels 0; Gaps 0;

OY 2 STELLVCLIRILRGAMSAKOHVGFEMAA 29  
 DB 203 STELLVCLIRILRGAMSAKOHVGFEMAA 230

RESULT 10  
 OY9B327 PRELIMINARY; PRT; 232 AA.  
 AC OY9B327;  
 DT 01-JUN-2001 (Tremblrel. 17, Created)  
 DT 01-JUN-2001 (Tremblrel. 17, Last sequence update)  
 DE CYTOCHROME OXIDASE SUBUNIT III (FRAGMENT).  
 GN COIII.  
 OS Peromyscus eremicus (cactus mouse).  
 OG Mitochondrion.  
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Sigmodontinae;  
 OC Peromyscus.  
 NCBI\_TaxID=42410;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RC STRAIN=2489\_TIBURON;  
 RA Hafner D.J., Riddle B.R., Alvarez-Castaneda S.T.;  
 RT "Evolutionary relationships of white-footed mice (Peromyscus) on  
 islands in the Sea of Cortez, Mexico."  
 RL Submitted (JAN-2001) to the EMBL/GenBank/DBJ databases.  
 DR EMBL: AF343760; AAK32172.1; -  
 KW Mitochondrion.  
 FT NON\_TER 1 1  
 FT 232 232  
 SQ SEQUENCE 232 AA; 26064 MW; FC1D518CAC4E6489 CRC64;

Query Match 52.8%; Score 75; DB 8; Length 232;  
 Best Local Similarity 53.6%; Pred. No. 0.00026;  
 Matches 15; Conservative 5; Mismatches 8; Indels 0; Gaps 0;

OY 2 STELLVCLIRILRGAMSAKOHVGFEMAA 29  
 DB 205 STELLVCLIRILRGAMSAKOHVGFEMAA 232

RESULT 11  
 OY9B321

ID OY9B321 PRELIMINARY; PRT; 232 AA.  
 AC OY9B321;  
 DT 01-JUN-2001 (Tremblrel. 17, Created)  
 DT 01-JUN-2001 (Tremblrel. 17, Last sequence update)  
 DE CYTOCHROME OXIDASE SUBUNIT III (FRAGMENT).  
 GN COIII.  
 OS Peromyscus collatus.  
 OG Mitochondrion.  
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Sigmodontinae;  
 OC Peromyscus.  
 NCBI\_TaxID=152343;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RC STRAIN=2437\_DATIL;  
 RA Hafner D.J., Riddle B.R., Alvarez-Castaneda S.T.;  
 RT "Evolutionary relationships of white-footed mice (Peromyscus) on  
 islands in the Sea of Cortez, Mexico."  
 RL Submitted (JAN-2001) to the EMBL/GenBank/DBJ databases.  
 DR EMBL: AF343768; AAK32180.1; -  
 KW Mitochondrion.  
 FT NON\_TER 1 1  
 FT 232 232  
 SQ SEQUENCE 232 AA; 26191 MW; A0491D58DA43A813 CRC64;

Query Match 52.8%; Score 75; DB 8; Length 232;  
 Best Local Similarity 53.6%; Pred. No. 0.00026;  
 Matches 15; Conservative 5; Mismatches 8; Indels 0; Gaps 0;

OY 2 STELLVCLIRILRGAMSAKOHVGFEMAA 29  
 DB 205 STELLVCLIRILRGAMSAKOHVGFEMAA 232

RESULT 12  
 OY9G3G9 PRELIMINARY; PRT; 233 AA.  
 AC OY9G3G9;  
 DT 01-MAR-2001 (Tremblrel. 16, Created)  
 DT 01-MAR-2001 (Tremblrel. 16, Last sequence update)  
 DE CYTOCHROME OXIDASE SUBUNIT III (FRAGMENT).  
 GN COIII.  
 OS Peromyscus boylii (brush mouse).  
 OG Mitochondrion.  
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Sigmodontinae;  
 OC Peromyscus.  
 NCBI\_TaxID=56316;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RC Pubmed=11083930;  
 RA Riddle B.R., Hafner D.J., Alexander L.F.;  
 RT "Phylogeography and Systematics of the Peromyscus eremicus Species  
 Group and the Historical Biogeography of North American Warm Regional  
 Deserts."  
 RL Mol. Phylogenet. Evol. 17:145-160(2000).  
 RN [2]  
 RP SEQUENCE FROM N.A.  
 RC Pubmed=11095731;  
 RA Riddle B.R., Hafner D.J., Alexander L.F., Jaeger J.R.;  
 RT "Cryptic vicariance in the historical assembly of a Baja California  
 Peninsular Desert biota."  
 RL Proc. Natl. Acad. Sci. U.S.A. 97:14438-14443(2000).  
 DR EMBL: AY009175; AAG45506.1; -  
 DR InterPro: IPR000298; CytoC\_oxdase\_III.  
 DR Pfam: PF00510; COX3; 1.  
 DR PROSITE: PS50253; COX3; 1.  
 KW Mitochondrion.  
 FT NON\_TER 1 1  
 FT 233 233





FT	NON_TER	233	233
SQ	SEQUENCE	233 AA;	26201 MW; E4154103AE4DE412 CRC64;

Query Match	52.88;	Score 75;	DB 8;	Length 233;
Best Local Similarity	53.68;	Pred. No. 0.00026;		
Matches 15;	Conservative 5;	Mismatches 8;	Indels 0;	Gaps 0;

QY 2 SIFLLVCLIRLRGAMSAKQHVGFEMAA 29  
 | | | | | : : | | | | |  
 Db 206 STFLMVCLLRQLKPHFTSKHHGFEMAA 233

Search completed: January 10, 2002, 02:04:05  
Job time: 217 sec

GenCore version 4.5  
Copyright (c) 1993 - 2000 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: January 10, 2002, 00:33:10 ; Search time 1538.44 Seconds  
(without alignments)  
607.682 Million cell updates/sec

Title: US-09-712-768-7

Perfect score: 87

Sequence: 1 gattcgatctctctgctgctggttccgagatgacgca 87

Scoring table:

IDENTITY\_NTC  
Gapop 10.0 , Gapext 1.0

Searched: 11351937 seqs, 5372889281 residues

Total number of hits satisfying chosen parameters: 22703874

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :

EST:\*  
1: em\_estfun:\*  
2: em\_esthum:\*  
3: em\_estlin:\*  
4: em\_estlom:\*  
5: em\_estlpl:\*  
6: em\_estlba:\*  
7: em\_estlro:\*  
8: em\_estlov:\*  
9: em\_hlc:\*  
10: gb\_estl:\*  
11: gb\_est2:\*  
12: gb\_hlc:\*  
13: gb\_gss:\*  
14: em\_gss\_fun:\*  
15: em\_gss\_hum:\*  
16: em\_gss\_inv:\*  
17: em\_gss\_pln:\*  
18: em\_gss\_pro:\*  
19: em\_gss\_rtd:\*  
20: em\_gss\_vtl:\*  
21: em\_gss\_other:\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

#### SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
C 1	30.6	35.2	700	10	AL509044
C 2	30.2	34.7	442	10	AA653844
C 3	30	34.5	440	10	BE592591
C 4	30	34.5	529	10	BE600048
C 5	29.8	34.3	271	11	BF152072
C 6	29.6	34.0	337	10	AA728883
C 7	29.6	34.0	409	10	AA227593
C 8	29.6	34.0	416	11	BF750333
C 9	29.6	34.0	457	10	AA446143
C 10	29.6	34.0	547	10	AA292266
C 11	29.2	33.6	1057	13	CNS00EFL
C 12	28.6	32.9	368	10	AW064184

13	28.6	32.9	659	13	AZ569835	AZ569835	267PVC04
14	28.6	32.9	663	13	AZ569886	AZ569886	267PVC03
C 15	28.4	32.6	397	10	AA192333	AA192333	z97g04.s
C 16	28.4	32.6	440	11	BF481009	BF481009	FM1_15-C0
C 17	28	32.2	238	10	AA627258	AA627258	ng64f01.s
C 18	28	32.2	433	10	AU090894	AU090894	AU090894
C 19	28	32.2	587	10	AM064799	AM064799	ST35H09.P
C 20	28	32.2	730	11	CR2311	CR2311	C82311
C 21	28	32.2	820	11	CR2068	CR2068	C82068
C 22	27.8	32.0	248	11	C14059	C14059	C14059
C 23	27.8	32.0	258	11	C14054	C14054	C14054
C 24	27.6	32.0	525	10	BE423205	BE423205	WHE0062.G
C 25	27.6	31.7	171	10	AA085035	AA085035	zn14c05.s
C 26	27.6	31.7	500	10	AU091009	AU091009	AU091009
C 27	27.6	31.7	565	10	BE234191	BE234191	140950.MA
C 28	27.4	31.5	360	10	AV202973	AV202973	AV202973
C 29	27.4	31.5	394	13	CNS020NM	AL206923	Tetraodon
C 30	27.4	31.5	565	10	AL502374	AL502374	AL502374
C 31	27.4	31.5	970	13	CNS02HLM	AL197789	Tetraodon
C 32	27.2	31.3	264	13	BA6969	BA6969	HS-1066-R2-
C 33	27.2	31.3	755	11	BP266774	BP266774	HV-CEA001
C 34	27.2	31.0	910	11	BI198381	BI198381	602761483
C 35	27.2	31.0	184	11	BG875316	BG875316	CM3-BT061
C 36	27	31.0	284	11	C14028	C14028	C14028
C 37	27	31.0	367	11	BF175955	BF175955	MY66132a
C 38	27	31.0	411	11	BG875632	BG875632	RCO-CN002
C 39	27	31.0	414	10	AA936801	AA936801	OH89B03.s
C 40	27	31.0	426	10	AA533464	AA533464	nj72d02.s
C 41	27	31.0	438	10	AV736142	AV736142	AV736142
C 42	27	31.0	440	10	AA845768	AA845768	ak77f05.s
C 43	27	31.0	457	10	A1073677	A1073677	ox98e05.s
C 44	27	31.0	475	10	AA635013	AA635013	ab47d11.r
C 45	27	31.0	496	10	AA707732	AA707732	zh23f09.s

#### ALIGNMENTS

RESULT 1  
AL509044/c 700 bp mRNA EST 04-JAN-2001  
LOCUS  
DEFINITION  
AL509044 Hordeum vulgare Barke developing carpypsis (3..15..DAP)  
ACCESSION  
AL509044 Hordeum vulgare cDNA clone HY10122V 5', mRNA sequence.  
VERSION  
AL509044.1 GI:12035547  
KEYWORDS  
EST.

#### SOURCE

ORGANISM

Hordeum vulgare  
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;  
Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae; Poideae  
; Triticeae; Hordeum.

1 (bases 1 to 700)

REFERENCE  
Michalek, W., Weschke, W., Pleissner, K., P. and Graner, A.

AUTHORS  
EST sequencing and analysis in barley

TITLE  
Unpublished (2000)

JOURNAL  
Contact: Michael W

COMMENT  
Institute for Plant Genetics and Crop Plant Research  
Corrensstr. 3, D-06466 Gatersleben, Germany  
Email: michael@ipk-gatersleben.de, http://pgrc.ipk-gatersleben.de  
Seq primer: T3 primer for 5' end.

FEATURES  
Location/Qualifiers

1..700  
/organism="Hordeum vulgare"  
/cultivar="Barke"  
/db\_xref="taxon:4513"  
/clone="HY10122V"  
/clone\_lib="Hordeum vulgare Barke developing carpypsis  
(3..15..DAP)"  
/tissue\_type="developing carpypsis (3..15..DAP)"  
/lab\_host="XLODR"  
/note="Vector: plasmid pBK-CMV; Site\_1: EcoRI; Site\_2:  
XhoI; mRNA was made from developing carpypsis (3..15..DAP)  
of spring barley variety 'Barke', a high quality malting



RESULT	4				
LOCUS	BE600048/c				
DEFINITION	BE600048	529 bp	mRNA	EST	18-AUG-2000
ACCESSION	PII_78-G06.b1	A002	Pathogen induced 1 (PII)	Sorghum bicolor cDNA,	
VERSION	BE600048				
KEYWORDS	BE600048.1	GI:9855121			
SOURCE	EST				
ORGANISM	sorghum.				
REFERENCE	Sorghum bicolor				
AUTHORS	Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta; Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae; PACC clade; Panicoideae; Andropogoneae; Sorghum.				
	1 (bases 1 to 529)				
	Cordonnier-Pratt,M.-M., Gingle,A., Dean,R., Sudman,M. and Pratt				
	J.L.H				
TITLE	An EST database from Sorghum: pathogen-induced plants				
JOURNAL	Unpublished (2000)				
COMMENT	Contact: Cordonnier-Pratt MM				
	Department of Botany				
	The University of Georgia				
	Plant Sciences Building, Rm. 2502, Athens, GA 30602-7271, USA				
	Tel: 706 542 1860				
	Fax: 706 542 1805				
	Email: mmpratt@uga.edu				
	Sequences have been trimmed to exclude POLYA, vector and regions				
	below phred quality 16. The threshold for highest quality sequences				
	is 20.				
	Seq primer: JEM REV				
	High quality sequence stop: 468				
	POLYA=No.				

```

FEATURES
source
Location/Qualifiers
1..529
/organism="Sorghum bicolor"
/db_xref="taxon:4556"
/clone_1lb="Pathogen induced 1 (P11)"
/note="Organ: Anthracnose-infected leaves from
two-week-old sorghum plants 48 hr after inoculation;
Vector: pBluescript II from lambda zap II; Site 1: XhoI;
Site 2: EcoRI; Two-week-old sorghum plants (BR7 632
cultivar) were infected with pathogen (isolate FR421 of
Colletotrichum graminicola, which is a sorghum isolate).
RNA was prepared from infected leaves harvested from 45
seedlings (2 weeks old) exhibit juvenile resistant
reaction, which is an incompatible interaction. As they
grow older (4 weeks or older), plants resume susceptibility
to anthracnose disease. The library was made from poly-A
RNA in the cloning vector lambda zap II. Clones to be
sequenced were prepared by mass excision. WARNING: While
most or all ESTs are expected to derive from the host
plant, no effort was made to eliminate ESTs deriving from
the pathogen."
71 a 204 c 146 g 108 t

```

[illegible]

DEFINITION	uz24c07.y1 NCI-CGAP Mamu Mus musculus cDNA clone IMAGE:366996 5' similar to SM:CA13 MOUSE P08121 COLLAGEN ALPHA 1(I)II CHAIN
ACCESSION	BT152072
VERSION	BT152072.1
KEYWORDS	EST.
SOURCE	house mouse.
ORGANISM	Mus musculus
REFERENCE	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
AUTHORS	NCI-CGAP <a href="http://www.ncbi.nlm.nih.gov/ncicgap">http://www.ncbi.nlm.nih.gov/ncicgap</a> .
TITLE	National Cancer Institute, Cancer Genome Anatomy Project (CGAP), Tumor Gene Index
JOURNAL	Unpublished (1997)
COMMENT	Contact: Robert Strausberg, Ph.D. Email: <a href="mailto:cgapbs-remail.nih.gov">cgapbs-remail.nih.gov</a> Tissue Procurement: Lothar Hennighausen Ph.D., Robin Humphreys cDNA Library Preparation: Life Technologies, Inc. cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LNU) DNA Sequencing by: Washington University Genome Sequencing Center Clone distribution: NCI-CGAP clone distribution information can be found through the I.M.A.G.E. Consortium/LNU at: <a href="http://image.llnl.gov/image/html/lresources.shtml">image.llnl.gov/image/html/lresources.shtml</a>

MGI:1430764

Trace considered overall poor quality

Seq primer: -40RP from Gibco

High quality sequence stop: 1.

Location/Qualifiers

1..271

						/organism="Mus musculus"		
						/strain="C57/B6"		
						/db_xref="taxon:10090"		
						/clone="IMAGE:366996"		
						/clone_1ib="NCI CGAP_Mam5"		
						/tissue_type="Tumor, gross tissue"		
						/dev_stage="7 months"		
						/lab_host="DH10B"		
						/note="Organ: mammary; Vector: PCW-SPORT6; site_1: SalI;		
						site_2: NotI; Cloned unidirectionally. Primer: Oligo dt.		
						library constructed by Life Technologies. Investigators		
						providing samples: Lothar Hennighausen/Robin Humphreys,		
						NH"		
BASE COUNT	34	a	77	c	80	g	80	t
ORIGIN								
Query Match							34.3%;	Score 29.8; DB 11; Length 271;
Best Local Similarity							63.0%;	Pred. No. 22;
Matches 46; Conservative							0;	Mismatches 27; Indels 0; Gaps 0;
Dy	11	tccgcgtgctgcgcgatccgcgcggcgagatgtcggcaaacagacgtcg	70					
Db	67	TCCGCTGGCGGCCCGCTGCATAGTGTCGTATGGTGATGTCGTGGACCCCTGGTGTCC	126					
Oy	71	gttcgagatgc	83					
Db	127	TGTTTCGCTGCC	139					

RESULT		6			
AA728883/c					
LOCUS					
DEFINITION	AA728883	337 bp	mRNA	EST	22-JAN-1998
	n337d11..1	NCl_CGAP-B5	Homo sapiens cDNA clone IMAGE:122389		
	similar to SW:P2C2_ARATH P49598 PROTEIN PHOSPHATASE 2C ; , mRNA sequence.				
ACCESSION	AA728883				
VERSION	AA728883.1	GI:2750242			
KEYWORDS	EST .				
SOURCE	human.				
ORGANISM	Homo sapiens				
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;					

REFERENCE 1 (bases 1 to 337)  
 AUTHORS NCI-CGAP <http://www.ncbi.nlm.nih.gov/ncicgap>.  
 TITLE National Cancer Institute, Cancer Genome Anatomy Project (CGAP),  
 JOURNAL Tumor Gene Index  
 COMMENT Unpublished (1997)  
 Contact: Robert Strausberg, Ph.D.  
 Email: cgaps-remail.nih.gov  
 Tissue Procurement: Ilan Kirsch, M.D., Kristina A. Cole, M.D.,  
 Ph.D. student, Rodrigo F. Chuquil, M.D., Michael R. Emmert-Buck,  
 M.D., Ph.D.  
 cDNA Library Preparation: David B. Krizman, Ph.D.  
 DNA Sequencing by: Washington University Genome Sequencing Center  
 Clone distribution: NCI-CGAP clone distribution information can be  
 found through the I.M.A.G.E. Consortium/BLN at:  
 www.bio.lnln.gov/dbirp/image/image.html  
 Insert length: 433 Std Error: 0.00  
 Seq primer: -28ml3 rev1 ET from Amersham.

FEATURES  
 source  
 1. 337  
 /organism="Homo sapiens"  
 /db\_xref="taxon:9606"  
 /clone\_lib="IMAGE:122389"  
 /clone\_lib="NCI-CGAP\_Br5"  
 /sex="female"  
 /tissue\_type="Infiltrating ductal carcinoma"  
 /lab\_host="DH10B"  
 /note="Organ: breast; Vector: pAMP10; mRNA made from  
 infiltrating ductal carcinoma, cDNA made by oligo-dT  
 priming. Non-directionally cloned. Size-selected on  
 agarose gel, average insert size 600 bp."

BASE COUNT 66 a 88 c 111 g 72 t  
 ORIGIN

Query Match 34.0%; Score 29.6; DB 10; Length 337;  
 Best Local Similarity 61.8%; Pred. No. 27;  
 Matches 47; Conservative 0; Mismatches 29; Indels 0; Gaps 0;

Qy 7 atcttcctgctgctcctgcgcatcctcgcggtgcgagtgcgcaaacagcac 66  
 Db 187 ATCTCTCTGTAGCTGCAGCTGTCACATCCAGCAGCTGTCGACAGAGACAA 128  
 Qy 67 gtccggttcgagatgg 82  
 Db 127 GAAGGCTTGAGCTGG 112

RESULT 7  
 AA227593 409 bp mRNA EST 06-AUG-1997  
 LOCUS z757409.r1 Soares\_NHMPU\_S1 Homo sapiens cDNA clone IMAGE:667480  
 DEFINITION 5', mRNA sequence.  
 ACCESSION AA227593  
 VERSION AA227593.1 GI:1849137  
 KEYWORDS EST.  
 SOURCE human.  
 ORGANISM Homo sapiens  
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 Mammalia; Eutheria; Primates; Catarrhini; Homiidae; Homo.  
 1 (bases 1 to 409)  
 Hillier, L., Allen, M., Bowles, L., Dubuque, T., Geisel, G., Jost, S.,  
 Kucaba, T., Lacy, M., Le, N., Lennon, G., Marra, M., Martin, J., Moore, B.,  
 Schellenberg, K., Steptoe, M., Tan, F., Theising, B., White, Y., Wyllie,  
 T., Waterston, R. and Wilson, R.  
 WashU-Merck EST Project 1997  
 UNPUBLISHED (1997)  
 CONTACT: Wilson RK  
 WASHINGTON UNIVERSITY School of Medicine  
 4444 Forest Park Parkway, Box 8501, St. Louis, MO 63108  
 Tel: 314 286 1800  
 Fax: 314 286 1810

EMAIL: est@watson.wustl.edu  
 This clone is available royalty-free through LNL; contact the  
 IMAGE Consortium ([info@image.lnln.gov](mailto:info@image.lnln.gov)) for further information.  
 Insert Length: 273 Std Error: 0.00  
 Seq primer: -28ml3 rev2 ET from Amersham.

FEATURES  
 source  
 1. 409  
 /organism="Homo sapiens"  
 /db\_xref="GDB:5561447"  
 /db\_xref="taxon:9606"  
 /clone\_lib="IMAGE:667480"  
 /clone\_lib="Soares\_NHMPU\_S1"  
 /tissue\_type="Pooled human melanocyte, fetal heart, and  
 pregnant uterus"  
 /lab\_host="DH10B"  
 /note="Organ: mixed (see below); Vector: pRT7D-Pac  
 (Pharmacia) with a modified polylinker; Site: 1; Not 1;  
 Site: 2; Eco RI; Equal amounts of plasmid DNA from three  
 normalized libraries (melanocyte 2NbHM, pregnant uterus  
 NBHPU, and fetal heart NBH19W) were mixed, and ss circles  
 were made in vitro. Following HAP purification, this DNA  
 was used as tracer in a subtractive hybridization  
 reaction. The driver was PCR-amplified cDNAs from pools of  
 5,000 clones made from the same 3 libraries. The pools  
 consisted of I.M.A.G.E. clones 260232-265223,  
 340488-345479, and 484488-489479."

BASE COUNT 95 a 104 c 121 g 89 t  
 ORIGIN

Query Match 34.0%; Score 29.6; DB 10; Length 409;  
 Best Local Similarity 61.8%; Pred. No. 28;  
 Matches 47; Conservative 0; Mismatches 29; Indels 0; Gaps 0;

Qy 7 atcttcctgctgctcctgcgcatcctcgcggtgcgagtgcgcaaacagcac 66  
 Db 83 ATCTCTCTGTAGCTGCAGCTGTCACATCCAGCAGCTGTCGACAGAGACAA 24  
 Qy 67 gtccggttcgagatgg 82  
 Db 23 GAAGGCTTGAGCTGG 8

RESULT 8  
 BF750333 416 bp mRNA EST 10-JAN-2001  
 LOCUS RCO-BN0416-261000-021-g06 BN0416 Homo sapiens cDNA, mRNA sequence.  
 DEFINITION BF750333  
 ACCESSION BF750333  
 VERSION BF750333.1 GI:12077009  
 KEYWORDS EST.  
 SOURCE human.  
 ORGANISM Homo sapiens  
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 Mammalia; Eutheria; Primates; Catarrhini; Homiidae; Homo.  
 1 (bases 1 to 416)  
 Dias Neto, E., Garcia Correa, R., Verjovski-Almeida, S., Briones, M.R.,  
 Nagai, M.A., da Silva, W. Jr., Zago, M.A., Bordini, S., Costa, F.F.,  
 Goldman, G.H., Carvalho, A.F., Matsukuma, A., Bata, G.S., Simpson, D.H.,  
 Brunstein, A., deoliveira, P.S., Bucher, P., Jongeneel, C.V., O'Hare,  
 M.J., Soares, F., Brentani, R.R., Reis, L.F., de Souza, S.J. and  
 Simpson, A.J.  
 Shotgun sequencing of the human transcriptome with ORF expressed  
 sequence tags  
 Proc. Natl. Acad. Sci. U.S.A. 97 (7), 3491-3496 (2000)  
 CONTACT: Simpson A.J.G.  
 Laboratory of Cancer Genetics  
 Ludwig Institute for Cancer Research  
 Rua Prof. Antonio Prudente 109, 4 andar, 01509-010, Sao Paulo-SP,  
 Brazil  
 Tel: +55-11-2704922  
 Fax: +55-11-2707001  
 Email: asimpson@ludwig.org.br







RESULT	15
AA192333/c	
LOCUS	AA192333
DEFINITION	AA192333.1 387 bp mRNA EST 16-JAN-1997
IMPGES:	zpr97g04.s1 Striatagene muscle 937209 Homo sapiens cDNA clone IMAGE:628182 3' similar to gb:U08441 CYTOCHROME C OXIDASE POLYPEPTIDE III (HUMAN), mRNA sequence.
ACCESSION	AA192333
VERSION	AA192333.1 GI:1781554
KEYWORDS	EST.
SOURCE	human.
ORGANISM	Homo sapiens
	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
	Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
REFERENCE	1 (bases 1 to 387)



## SUMMARIES

Result	No.	Score	Query Match	Length	DB	ID	Description
	1	87	100.0	87	6	AX113992	AX113992 Sequence
	2	48.4	55.6	286550	1	SME591785	AL591785 Sinorhizo
	3	46.8	53.8	4890	1	PDCOX1	X05882 Paracoccus
	4	45.4	52.2	10835	1	AE008010	AE008010 Agrobacte
	5	43.6	50.1	12299	1	AE006000	AE006000 Caulobact
	6	43.6	50.1	346510	1	AE003011	AE003011 Mesorhizo
	7	42	48.3	8121	1	BUA242392	BUA242392 Bradyrhiz
	8	42	48.3	31495	1	BJ033883	AE004449 Pseudomon
	9	35.2	40.5	12829	1	AE004449	AE004449 Pseudomon
	10	31	35.6	294800	1	SME591789	AL591789 Sinorhizo
	11	30.4	34.9	227	8	AY023455	AY023455 Oryza sat
	12	30.2	34.7	714	9	MACC903A	M85148 Macaca mula
	13	30	34.5	176383	9	AC019055	AC019055 Homo sapi
	14	29.8	34.3	125110	9	AL450267	AL450267 Human DNA
	15	29.8	34.3	185147	2	AP002388	AP002388 Homo sapi
	16	29.2	33.6	83463	9	AL161644	AL161644 Human DNA
	17	29.2	33.6	151552	2	AL357673	AL357673 Homo sapi
	18	28.6	32.9	16472	9	HIMT5SEQ	X99256 Hylobates l
	19	28.4	32.6	1325	1	MBU62766	U62766 Mycobacteri
	20	28.4	32.6	110000	2	LMFCHNR36_02	Continuation (3 of
	21	28.4	32.6	110000	2	LMFCHNR36_03	Continuation (4 of
	22	28.4	32.6	209175	2	AL450331	AL450321 Mus muscu
	23	28.2	32.4	5642	1	AF228578	AF228578 Rhizobium
	24	28	32.2	690	10	AY010181	AY010181 Ammospert
	25	28	32.2	690	10	AY010184	AY010184 Ammospert
	26	28	32.2	690	10	AY010185	AY010185 Ammospert
	27	28	32.2	690	10	AY010186	AY010186 Ammospert
	28	28	32.2	690	10	AY010194	AY010194 Ammospert
	29	27.8	32.0	2147	5	AB005533	AB005533 Coturnix
	30	27.6	31.7	1591	3	AF397532	AF397532 Drosophila
	31	27.6	31.7	1604	3	AF218237	AF218237 Drosophila
	32	27.6	31.7	9429	2	AC020441	AC020441 Drosophila
	33	27.6	31.7	16884	4	TE0Y19192	Y19192 Talpa europ
	34	27.6	31.7	37750	2	LMFCHNR28_0	LMFCHNR28_0
	35	27.6	31.7	110000	2	LMFCHNR32_05	LMFCHNR32_05
	36	27.6	31.7	156508	3	AC007581	AC007581 Drosophila
	37	27.6	31.7	170089	3	AC007925	AC007925 Drosophila
	38	27.6	31.7	189370	1	AF010496	AF010496 Rhodobact
	39	27.6	31.7	194815	9	CNS010VZ	AL136295 Human chr
	40	27.6	31.7	328500	3	AE003465	AE003465 Drosophila
	41	27.4	31.5	2814	1	AF315510	AF315510 Lysobacte
	42	27.4	31.5	27644	3	CEIK0282	Y14558 Geobacterb
	43	27.4	31.0	12363	1	AE002088	AE002088 Deinococc
	44	27	31.0	16561	9	PMWITG	X93335 Pan troglod
	45	27	31.0	16561	9	PMWITG	X93335 Pan troglod

## PAT 01-MAY-2001

RESULT	1				
AXI13992					
LOCUS	AXI13992	87 bp	DNA	PAT	01-MAY-2001
DEFINITION	Sequence 7 from Patent EP1105603.				
ACCESSION	AXI13992				
VERSION	AXI13992.1	GI:13940152			
KEYWORDS					
SOURCE	Gluconobacter oxydans.				
ORGANISM	Gluconobacter oxydans				
	Bacteria; Proteobacteria; alpha subdivision; Acetobacteraceae				
	Gluconobacter.				

REFERENCE  
AUTHORS  
TITLE  
JOURNAL

1 (bases 1 to 87)  
Asakura, A., Hoshino, T. and Shinjoh, M.  
Cytochrome c oxidase complex from gluconobacter oxydans  
Patent: EP 1103603-A 7 30-MAY-2001;  
EPO 1103603-A 7 30-MAY-2001;

## FEATURES

```
/organism="Gluconobacter oxydans"  
/db_xref="taxon:442"
```

CDS	/note=>87 /name="unnamed protein product" /codon_start=1 /transl_table=11 /protein_id="CAC37872.1" /db_xref="GI:13940153" /translation="DSIFLVLCVIRILRGAMSAKOHVFEMAA"
BASE COUNT	14 a 26 c 26 g 21 t
ORIGIN	
Query Match	100.0%; Score 87; DB 6; Length 87;
Best Local Similarity	100.0%; Pred. No. 2.2e-14;
Matches 87; Conservative 0; Mismatches 0; Indels 0; Gaps 0;	
Db	1 GATTGATCTTCGGTGGTGCTGTGCCTGATCCGACATCCTCGCGGTGCATGTCTGCCAATA 60
QY	61 CAGCAGTCGGTTTCGAGATGCGCA 87
Db	61 CAGCAGTCGGTTTCGAGATGCGCA 87
RESULT 2	SME591785 286550 bp DNA BCT 16-AUG-2001
LOCUS	SME591785 286550 bp DNA BCT 16-AUG-2001
DEFINITION	Sinorhizobium meliloti 1021 complete chromosome; segment 4/12.
ACCESSION	AL591785 AL591688
VERSION	AL591785.1 GI:15073729
KEYWORDS	Sinorhizobium meliloti. Sinorhizobium meliloti. Bacteria; Proteobacteria; alpha subdivision; Rhizobiaceae group; Rhizobiaceae; Sinorhizobium.
SOURCE	1 (bases 1 to 286550) Capela,D., Barloy-Hubler,F., Gouzy,J., Bothé,G., Ampe,F., Batut,J., Boistard,P., Becker,A., Boutry,M., Cadieu,E., Dreano,S., Gloux,S., Gottie,T., Goffeau,A., Kahn,D., Kiss,B., Lelaure,V., Masuy,D., Pohl,T., Portetelle,D., Puehler,A., Putnelle,B., Ramsperger,U., Renard,C., Thebaud,P., Vandenbol,M., Weidner,S. and Galibert,F. From the Cover: Analysis of the chromosome sequence of the legume symbiont Sinorhizobium meliloti strain 1021 Proceedings of the National Academy of Sciences of the United States of America. 98 (17), 9877-9882 (2001)
JOURNAL	11481430 2 (bases 1 to 286550) Gouzy,J. Direct Submission Submitted (26-JUL-2001) Gouzy J., Submitted on behalf of the MELILO EU Consortium
PUBMED	MELILO EU Consortium: Laboratoire de Biologie Moléculaire des Relations Plantes-Microorganismes, UMR215-CNRS-INRA, BP27, F-31326 Castanet, France, Laboratoire de Génétique et Développement UMR6061-CNRS, Faculté de Médecine, 2 avenue du Pr. Leon Bernard, F-35043 Rennes, France, GATC GmbH, Fritz-Arnold-str. 23, D-78467 Konstanz, Germany, Universitaet Bielefeld, Biologie IV (Genetik) Universitaetstr 25, D-33615 Bielefeld, Germany, Unite de Biochimie physiologique, Université Catholique de Louvain, Place Croix du Sud 2, Bte 20, B-1348 Louvain-la-Neuve, Belgium, Unite de Microbiologie, Faculté des Sciences Agronomiques de Gembloux, Avenue Mercchal Juin 6, B-5030 Gembloux, Belgium. E-mail:jerome.gouzy@loulouise.inra.fr <a href="http://sequence.loulouse.inra.fr/meliloti.html">http://sequence.loulouse.inra.fr/meliloti.html</a> .
REFERENCE	1. 286550 Location/Qualifiers /organism="Sinorhizobium meliloti" /strain="1021" /db_xref="taxon:382" 70..813 /gene="gstII OR SMC00916" 70..813 /gene="gstII OR SMC00916"
AUTHORS	gene CDS
TITLE	repeat_region
JOURNAL	gene CDS
PUBMED	gene CDS
DEFINITION	gene CDS
ACCESSION	gene CDS
VERSION	gene CDS
KEYWORDS	gene CDS
SOURCE	gene CDS
SOURCE	gene CDS
DEFINITION	gene CDS
ACCESSION	gene CDS
VERSION	gene CDS
KEYWORDS	gene CDS
SOURCE	gene CDS
SOURCE	gene CDS
DEFINITION	gene CDS
ACCESSION	gene CDS
VERSION	gene CDS
KEYWORDS	gene CDS
SOURCE	gene CDS
SOURCE	gene CDS
DEFINITION	gene CDS
ACCESSION	gene CDS
VERSION	gene CDS
KEYWORDS	gene CDS
SOURCE	gene CDS
SOURCE	gene CDS
DEFINITION	gene CDS
ACCESSION	gene CDS
VERSION	gene CDS
KEYWORDS	gene CDS
SOURCE	gene CDS
SOURCE	gene CDS
DEFINITION	gene CDS
ACCESSION	gene CDS
VERSION	gene CDS
KEYWORDS	gene CDS
SOURCE	gene CDS
SOURCE	gene CDS
DEFINITION	gene CDS
ACCESSION	gene CDS
VERSION	gene CDS
KEYWORDS	gene CDS
SOURCE	gene CDS
SOURCE	gene CDS
DEFINITION	gene CDS
ACCESSION	gene CDS
VERSION	gene CDS
KEYWORDS	gene CDS
SOURCE	gene CDS
SOURCE	gene CDS
DEFINITION	gene CDS
ACCESSION	gene CDS
VERSION	gene CDS
KEYWORDS	gene CDS
SOURCE	gene CDS
SOURCE	gene CDS
DEFINITION	gene CDS
ACCESSION	gene CDS
VERSION	gene CDS
KEYWORDS	gene CDS
SOURCE	gene CDS
SOURCE	gene CDS
DEFINITION	gene CDS
ACCESSION	gene CDS
VERSION	gene CDS
KEYWORDS	gene CDS
SOURCE	gene CDS
SOURCE	gene CDS
DEFINITION	gene CDS
ACCESSION	gene CDS
VERSION	gene CDS
KEYWORDS	gene CDS
SOURCE	gene CDS
SOURCE	gene CDS
DEFINITION	gene CDS
ACCESSION	gene CDS
VERSION	gene CDS
KEYWORDS	gene CDS
SOURCE	gene CDS
SOURCE	gene CDS
DEFINITION	gene CDS
ACCESSION	gene CDS
VERSION	gene CDS
KEYWORDS	gene CDS
SOURCE	gene CDS
SOURCE	gene CDS
DEFINITION	gene CDS
ACCESSION	gene CDS
VERSION	gene CDS
KEYWORDS	gene CDS
SOURCE	gene CDS
SOURCE	gene CDS
DEFINITION	gene CDS
ACCESSION	gene CDS
VERSION	gene CDS
KEYWORDS	gene CDS
SOURCE	gene CDS
SOURCE	gene CDS
DEFINITION	gene CDS
ACCESSION	gene CDS
VERSION	gene CDS
KEYWORDS	gene CDS
SOURCE	gene CDS
SOURCE	gene CDS
DEFINITION	gene CDS
ACCESSION	gene CDS
VERSION	gene CDS
KEYWORDS	gene CDS
SOURCE	gene CDS
SOURCE	gene CDS
DEFINITION	gene CDS
ACCESSION	gene CDS
VERSION	gene CDS
KEYWORDS	gene CDS
SOURCE	gene CDS
SOURCE	gene CDS
DEFINITION	gene CDS
ACCESSION	gene CDS
VERSION	gene CDS
KEYWORDS	gene CDS
SOURCE	gene CDS
SOURCE	gene CDS
DEFINITION	gene CDS
ACCESSION	gene CDS
VERSION	gene CDS
KEYWORDS	gene CDS
SOURCE	gene CDS

```

/codon_start=1
/transl_table=11
/evidence=not_experimental
/product="60 KD CHAPERONIN A PROTEIN"
/protein_id="CAC45364.1"
/db_xref="GI:15073723"
/translation="MAAKEVKEGSRAREKMLRGVDIADAVKVTLGPRGNVVDKSF
GAPRTTKDGVSAKEIELEDKFEENKGAQMREVASKTNIDAGDGTATVLAQAIIVE
GAKAAAGMNPMDLKGIDLAIAEVNDLAKARKINTSDPAQVGTISANGEQITGL
DIAEMOKVNEGVTVEAKTAETAELEEVGMPDGYISPTVTVNPKRWADLEDA
FILHEKLSNLMQMLPYLEAVVOTGKFLITADVESEALATLVNKGKIAAV
KAPGDRKRMALEDIAITGTVISBDLGKLSVYLDMDGKRAKVSITKENTIID
GAGKSDIEGRVAQIKAIETETSDYDREKLOERLAKAGVAIRVGATEVEVK
KDRIDALNATRAVQEGIVGCGVALIRSSVKITVKGENDODAGVNIYRALQSPA
ROIYENAGDEASIVVKGITKEKNDPFGYNQOTGEGMIMGIIDIPKAVVTALQDA
SVASLTITTEMIMELPKKAPAMPGMGMGGMGNDMA"
complement(4298..4594)
/gene="groES1 OR SMC00912"
complement(4298..4594)
/gene="groES1 OR SMC00912"
/function="cell processes: chaperoning"
/notes="predicted by Codon_usage
predicted by Homology
predicted by Framed"
/codon_start=1
/transl_table=11
/evidence=not_experimental
/product="10 KD CHAPERONIN A PROTEIN"
/protein_id="CAC45365.1"
/db_xref="GI:15073724"
/translation="MASTNRPRLHDRVVRVSESEKTKGGIIPDTAKERPOEGEIV
AVSGAGDESGKVPILDKAGDRILFGKSGTEVKINGEDLLMKKEADINGVIG"
4863..5690
/gene="SMC00911"
4863..5690
/gene="SMC00911"
/function="miscellaneous: hypothetical/global homology"
/notes="Product confidence: hypothetical
Gene name confidence: hypothetical
predicted by Codon_usage
predicted by Framed"
/codon_start=1
/transl_table=11
/evidence=not_experimental
/product="CONSERVED HYPOTHETICAL PROTEIN"
/protein_id="CAC45366.1"
/db_xref="GI:15073725"
/translation="MASFNSISSEKLARLVGTRKAPVYIDVRDDDFADHILVPSF
HRGHVHPQAEYRGWVYVLCRHGKLSGSAAMLRHGAAGAEIILDGIEAMRAG
LPLVPASLPNSDPAGSVMVTRSRKIDRIACPMILRRVDPARFLVTPAEVAV
GERRGATPFDIEGVFWSHREGECTFDTMVEGLSFPALPHVARIYRGADTREREIP
EAAGLAVSISGLSRMYADDLQLEAGMTIYDALYRWARADATQEKHDVSHSGKGRKP"
6004..6852
/gene="SMC00910"
6004..6852
/gene="SMC00910"
/function="miscellaneous: hypothetical/global homology"
/notes="Product confidence: hypothetical
Gene name confidence: hypothetical
predicted by Codon_usage
predicted by Homology
predicted by Framed"
/codon_start=1
/transl_table=11
/evidence=not_experimental
/product="CONSERVED HYPOTHETICAL PROTEIN"
/protein_id="CAC45367.1"
/db_xref="GI:15073726"
/translation="MAVRINSEREIAGRYDVVLCWGVILHNGIQAFASACBALAAR
AAGLTIVLITNSRPHGVYVQIRGLGVDEAVDRIVTSGDVQALIAADRIRIFIG
ADRDLPLEIGITEIVASDAEETIVCAFYDDEETPEHYRATVIGLARKIIPICAN
PDLVVERGHRILPCAGAIKAIKIEELGEGARIGKPKYIATYRAALAEAKARAFQISR
VIAIGDMPTDVKAGADAGFDLLIYSAGIHAEYMHESITDARIMAFILKQNGAOPKW

```

---

```

gene      MPRLA"
6866..7849
/gene="ribD OR SMC00909"
6866..7849
/gene="ribD OR SMC00909"
/EC_number="2.7.1.26"
/EC_number="2.7.1.2"

Query Match      55.68; Score 48.4; DB 1; Length 286550;
Best Local Similarity 74.4%; Pred. No. 0.00036;
Matches 61; Conservative 0; Mismatches 21; Indels 0; Gaps 0;

Db 130642 CGATCTTCCTGCTGCTGCTGCTGCGCGCCCTCGCGGCGACTTCACCGCAAGCAC 130701
QY 65 acgtcggttcgagatgagccgc 86
|| ||| ||||| ||| ||
Db 130702 ACTTCGGCTTCGAGCGCGCGC 130723

RESULT 3
LOCUS PDCOX1 4890 bp DNA BCT 12-SEP-1993
DEFINITION Paracoccus denitrificans cytochrome c oxidase CoII and CoIII genes
ACCESSION X05828
VERSION X05828.1 GI:45468
KEYWORDS CoII gene; cytochrome aaz; cytochrome c oxidase; unidentified reading frame.
SOURCE Paracoccus denitrificans.
ORGANISM Paracoccus denitrificans.
Bacteria; Proteobacteria; alpha subdivision; Rhodobacter group; Paracoccus.
REFERENCE 1 (bases 1 to 4887)
AUTHORS Raitio,M., Jalli,T. and Saraste,M.
TITLE Isolation and of the genes for cytochrome c oxidase in Paracoccus denitrificans
JOURNAL EMBO J. 6, 2825-2833 (1987)
REFERENCE 2 (bases 1 to 4890)
AUTHORS Saraste,M.
TITLE Direct Submission
COMMENT Submitted (07-MAR-1988) to the EMBL/Genbank/DBJ databases
FEATURES
source
1..4890
/organism="Paracoccus denitrificans"
/db_xref="taxon:266"
/map="cox1 locus"
853..856
/notes="gccc was gcg in [1]"
/citation=[1]
862..863
/notes="ag was a in [1]"
/citation=[1]
880..882
/notes="agt was at in [1]"
/citation=[1]
892..894
/notes="tgg was tg in [1]"
/citation=[1]
898..900
/notes="gcc was gc in [1]"
/citation=[1]
900..906
/notes="inverted repeat A"
912..915
/notes="TATA-box like sequence"
924
/notes="g was c in [1]"
/citation=[1]
937..938
/notes="tg was tcgc in [1]"

```

```

repeat_unit
940..946
/note="inverted repeat A'"
984..987
/note="pot. ribosome binding site"
992..1888
/note="coil gene product (AA 1 - 298)"
/codon_start=1
/transl_table=11
/protein_id="CAA29268.1"
/db_xref="GI:45469"
/translation="MMAIATKRGVAAVMSIGVATMTAPVALADVLGDLPVIGKPVN
GGMSPASSPLPHDOOQMLDHFVLYITPAVITFPCLLILITVFNRRANVPAPKFRH
NPPIVMTIVTILVAVIGAFSLIPIDRSEOMENDPDLVTKALGHOMVSYEYNDGQ
VAFDMLTEKRALDNAGVSEDEYLLATDNPVYVQGVKVLQVATVDYIHAMTIPAPA
VKQDAVPGGRHQLMFSVDQEGVYFEQCSELCGINHAIKPIYKKAVSQEKTEAMLAGAK
EEFADASDYLPSAPVKLIASAE"
1961..1965
/note="pot. ribosome binding site"
1974..2813
/note="ORF 1 (AA 1 - 279)"
/codon_start=1
/transl_table=11
/protein_id="CAA29269.1"
/db_xref="GI:45470"
/db_xref="SWISS-PROT:P08301"
/translation="MSLVFTAFVGLMTAPQVPNPEVAFCAVLFIATGGASGALNMNM
VDADIDAWRRTRAGRPSPGSRVTSQEPALPAGVIALSGLSVMLGAGWMEFAGFLPITI
PYAVVYVITLKRSTPQNIIVIGAGAPAPMIGNALPFGIGTISLMAFLIFWPPR
HEWALALEMKDYSKAGVPMILVTYHGRKVTYCHIFATITLVAPPALMLGTFSVCGPLX
LAVSVALNFIATGGWQILIRSEDDQADGYRVEKRIFLSLTYTFILFALLVQHWV
GGM"
2802..2806
/note="pot. ribosome binding site"
2813..3082
/note="ORF 2 (AA 1 - 89)"
/codon_start=1
/transl_table=11
/protein_id="CAA29270.1"
/db_xref="GI:45471"
/db_xref="SWISS-PROT:P08302"
/translation="MLPVEHEHLKRRSRNIGLIVLLAFVALVEGLSVYKIIQGDH
MGGYDHRPASMLPPDPDPAPNAPVAPAPGPRATQSGVTOEANG"
3079..3663
/note="ORF 3 (AA 1 - 194)"
/codon_start=1
/transl_table=11
/protein_id="CAA29271.1"
/db_xref="GI:45472"
/db_xref="SWISS-PROT:P08303"
/translation="MSGGRKRSNRTVAMLAGVYVLMGALSMAAVPIFSWCKYTGFR
GTPTVAEASDVIDLEKIKRFPDANADSNLQTERPMQREMLKIGENALFYELININ
GEBVYFQSYVWAPDAAGVFPNKEICFCFTQETLQDPERVEMPSPFVADLVNDRI
AYIRDTLTSYFHRDPPAPKQAAALDAKTEPIVN"
3666..3691
/note="pot. ribosome binding site"
3698..4522
/note="coil gene product (AA 1 - 274)"
/codon_start=1
/transl_table=11
/protein_id="CAA29272.1"
/db_xref="GI:45473"
/db_xref="SWISS-PROT:P06030"
/translation="MAHYKNHXYOILPSTIMPFCAIGAFVMLTGAVANMKGIPEEG
IPEBGPMELILIGLVLYVMEGMADVYVNEGTEGRTPYVRIGLQYGLTILINSEMPF
VAEVMFENINALYPMGPDSPIKGVWPBGIIVTEPWHPLPLNTILILLSGVAFTWAL
HAFVMEIGDKRTKTINGLIVAVILGVCFGTQDAYEYSHAAGLADIVYAAAFWMAEGFH
AHVITICITLFCOLLRLKGMQTKOROHGFEFAAAYVMEVDVWVLEFVVIYIMGR"
4543..4549
/note="inverted repeat B"
4543..4574
/note="pot. transcription terminator"

```

BASE COUNT	777 a	1581 c	1623 g	909 t	ORIGIN
Query Match	53.8%;	Score 46.8;	DB 1;	Length 4890;	
Best Local Similarity	73.2%;	Pred. No. 0.002;			
Matches 60;	Conservative 0;	Mismatches 22;	Indels 0;	Gaps 0;	
Db 4431	AGTGGCTTCGAGGCCCGCCG	4452			
RESULT 4	AE008010	10835 bp	DNA	BC1	14-AUG-2001
LOCUS	AE008010	10835 bp	DNA	BC1	14-AUG-2001
DEFINITION	Agrobacterium tumefaciens strain C58 circular chromosome, section 68 of 254 of the complete sequence.				
ACCESSION	AE008010	AE007869			
VERSION	AE008010.1	GI:1515740			
KEYWORDS	Agrobacterium tumefaciens.				
SOURCE	Agrobacterium tumefaciens				
ORGANISM	Bacteria; Proteobacteria; alpha subdivision; Rhizobiaceae group; Rhizobiaceae; Rhizobium.				
REFERENCE	1 (bases 1 to 10835)				
AUTHORS	Hinkle,G., Slater,S.C. and Goodner,B.				
TITLE	Complete Genome Sequence of Agrobacterium tumefaciens C58 (Rhizobium radiobacter C58), the Causative Agent of Crown Gall Disease in Plants				
JOURNAL	Unpublished				
REFERENCE	2 (bases 1 to 10835)				
AUTHORS	Hinkle,G., Slater,S.C. and Goodner,B.				
TITLE	Direct Submission				
JOURNAL	Submitted (14-AUG-2001) Cereon Genomics, 45 Sidney Street, Cambridge, MA 02139, USA				
FEATURES	Location/Qualifiers				
Source	1..10835				
gene	/organism="Agrobacterium tumefaciens"				
gene	/strain="C58"				
gene	/db_xref="taxon:358"				
gene	complement(118..642)				
gene	/gene="AGR_C.1396"				
gene	complement(118..642)				
gene	/gene="AGR_C.1396"				
gene	/note="(AJ242592) hypothetical protein"				
gene	/codon_start=1				
gene	/transl_table=1				
gene	/product="AGR_C.1396p"				
gene	/protein_id="AAK6576.1"				
gene	/db_xref="GI:1515741"				
gene	/translation="MRSLPLAKFTLLAAGLAIIIPAAALAAQDPPECTPKSTHGAMSVICDPVAGASEQCALMOWNIADDPREGLSVITKTADRSRTITLAPLGLLKDGMELVLDNNNIGRAYFTRCPSSEGCYVEVDIDELKLVLAKGNVAFALRESVDQDRVGILPILSLGAFAGYDALP"				
gene	988..1938				
gene	/gene="AGR_C.1397"				
gene	988..1938				
gene	/gene="AGR_C.1397"				
gene	/note="(AJ242592) cytochrome c oxidase subunit II"				
gene	/codon_start=1				
gene	/transl_table=1				
gene	/product="AGR_C.1397p"				
gene	/protein_id="AAK6576.1"				
gene	/db_xref="GI:1515742"				
gene	/translation="MDSASVTDLDQTRGECTYVNNRIYALAGTCLLTAGAHDQPVHWMQGEAATPIIMHETRMFEQYTLWFIIVPTLFTVLLALITIVAKHAAKNPVASTV				

gene	SHNTAIEVWTLAPLILFLAPSPFNLLNAOLTOPENPDLLKATATOLMSLEYKA AEGAPLSFSDYLLKDDODRAAEKEDKARYBRLLAVDNEMVVPYKTVRLLYTAAPD VIAHAPAMPARGVKIDAVPGRLENTWPFPEKRGILTYGCCSEICGKHAFMPLAIRVSE QOYTWHAANAASDINGANRALMASVDGAPRTVDVANAETN"	/codon_start=1 /transl_table=11 /product="AGR_C_1409p" /protein_id="AAK86581.1" /db_xref="GI:15155747"
CDS	2549..3664 /gene="AGR_C_1399" /note="CYTOCHROME C OXIDASE POLYPEPTIDE I (CYTOCHROME AA3 SUBUNIT 1)"	/codon_start=1 /transl_table=11 /product="AGR_C_1399p" /protein_id="AAK86577.1" /db_xref="GI:15155743"
gene	SVYVTAFLLSPLVLAGITMLTDNFTAFSPGSGDPILLYOHLFWFQHPVEY ILILPGFIISHIVSTPSKRPVGLGMAYAMAIGAVGIWNAHMYIVGLSLEOR YFVATVIAVPTGRIKIFSWIATWVGSLFSTPMWVAIGFIFLETYGVGYOLANA GLDSLDHYVVAHFHYVLSLGAVFIFAGWYWPCKITGYWNEFIGKLFHFWMI GVNLIFPPHGLIAGMPRRYIDPDYAGMNVVSSYGSYISAVAGIPLFGWENAF KKRIAGNPPGCEGATILENQLSSPPPHOQLPRIR"	/transl_table=11 /product="AGR_C_1402p" /protein_id="AAK86578.1" /db_xref="GI:15155744"
CDS	4054..5007 /gene="AGR_C_1402" /note="(X89566) heme O synthase"	/codon_start=1 /transl_table=11 /product="AGR_C_1402p" /protein_id="AAK86578.1" /db_xref="GI:15155744"
gene	LVLAPGRINPLIGLITICTAGAGASGAINMYDADIDVMTTRAKRPLPSRIAPR EALAFGLTASPSVITIGLVNPFSGCLFTTFPAVYTMWKSTPONTIVIGAA GAPRPLGMACVGVGSLSVLI.FLITFMTPAHFALALFKMDGAVGIDPMPVVA GERSTKNQMTIVAVLTAANAAPFETGLSAGYICPAAYLSAIFVYCSFDRMRPBGD EKMLPAKMFAYSVLYLFAIFSLGLADHFAPALKAVISGL"	/transl_table=11 /product="AGR_C_1406p" /protein_id="AAK86579.1" /db_xref="GI:15155745"
CDS	5167..5784 /gene="AGR_C_1406" /note="CYTOCHROME C OXIDASE ASSEMBLY PROTEIN CTAG"	/codon_start=1 /transl_table=11 /product="AGR_C_1406p" /protein_id="AAK86579.1" /db_xref="GI:15155745"
gene	LFCRVYNGTTOREYSDVILDKTINVFEDNTSGINMDPRPYDKMKPKIGETI KYTFKATNRSPVATGTAAFNVTPEMGAAYFNKVEGCFEETTLQPGFTLEMPVFEV DPDIATARETKSVHTLSTFTYTPAKTERPVASLPVKTESGSKL"	/transl_table=11 /product="AGR_C_1407p" /protein_id="AAK86580.1" /db_xref="GI:15155746"
CDS	5846..6721 /gene="AGR_C_1407" /note="(A0242592) cytochrome c oxidase subunit III"	/codon_start=1 /transl_table=11 /product="AGR_C_1407p" /protein_id="AAK86580.1" /db_xref="GI:15155746"
gene	ASFVMEFVAMFAYPDASLPFPHATGASLETTGCGWPKGLEVIDPMLPLNTVIL LLSGTCVTAHNAHLNDRKGLISGLATLVALGVLPSTVOVEYIHAAPDFNNSIYGA TFPMATGHEGFHFVGVTFVLVCLFRAIAGGFTPKOHGFEEAAMWYHVDVWMLFLF FALYIWGGWGAPEH"	/transl_table=11 /product="AGR_C_1417p" /protein_id="AAK86582.1" /db_xref="GI:15155748"
CDS	6941..7369 /gene="AGR_C_1409" /note="(A0242592) hypothetical protein"	/codon_start=1 /transl_table=11 /product="AGR_C_1409p" /protein_id="AAK86582.1" /db_xref="GI:15155748"
gene	THQGGTYITYPELTLADGRLLFNKGFVPEPMKDAARPDGVSGETYITGLAAPL VAKPSSLPHDNDIAKNIFFWKDLAAMASSAEVPEPLVKLFVDADNAPNPGWPGGV TLIDLPHNLQYAITWYGLAALVIVAGFAYFRNKGAGE"	/transl_table=11 /product="AGR_C_1414p" /protein_id="AAK86583.1" /db_xref="GI:15155749"
CDS	8173..9219 /gene="AGR_C_1414" /note="(A006203) LytB"	/codon_start=1 /transl_table=11 /product="AGR_C_1414p" /protein_id="AAK86583.1" /db_xref="GI:15155749"
gene	REHIEYHNRVYVGLBAKGAIEVEEHEIIPAEHREQVYFSAHGVKSPEDAOANLF YDACEPLNLSKVKQAMRHRORIGRHVILIGHAGHEPVTGNGOLPEGVSVIYEDV GYEPEVDRNCEFTVOTTLSDVDTAGVILARQEREPALQAPPAADSICATINRDVAY QAAKCDLEIVYVAGRNSSSKRLVYVALRAQKHSVLYVORASEIDMNTGDIRVGLS AGASAPDEVDELIEAFKARPDITDLAVVEETHEFLVNLRLSIELTDDMAFVNG NASNALPRAKTAGI"	/transl_table=11 /product="AGR_C_1416p" /protein_id="AAK86584.1" /db_xref="GI:15155750"
CDS	9312..10280 /gene="AGR_C_1416" /note="HOMOSERINE KINASE (HK)"	/codon_start=1 /transl_table=11 /product="AGR_C_1416p" /protein_id="AAK86584.1" /db_xref="GI:15155750"
gene	KDPLITLTYEKREKENDLPFLGLMOMLAAGLSCLPLPKRDELLOGLSGRAALI SFLGCMWLKPEAKHCRREVGAALAMHLASGFEIKRNALSYDKMVLMDKSEERAD EVEKGLREIREIRPIDYLAHMPKDLPAVIAHADLEQDMPFLGDELGLIDFYRACND LLAYVSIICINAMCPKEDGAYNWKFGKALLGEGYSVPLSRAELBAIPLSRGSAIRP FLTRLYDMLTYPAGALVYKRPDLETLRKLRHRTIANAEYGLAGE"	/transl_table=11 /product="AGR_C_1417p" /protein_id="AAK86585.1" /db_xref="GI:15155751"
CDS	10277..10717 /gene="AGR_C_1417" /note="Ribonuclease H (E.C.3.1.26.4) Mutant with His 62 Replaced by Pro (H62P)"	/codon_start=1 /transl_table=11 /product="AGR_C_1417p" /protein_id="AAK86585.1" /db_xref="GI:15155751"
gene	RHEILAIASALNAKSPCEVDLYTDSATYKDGITKWTFGWKKKGKTPADNNPVKNVEL WQALEAQRERKVTILHWYKGHAGHPENRADELARKGMEPKRR"	/transl_table=11 /product="AGR_C_1417p" /protein_id="AAK86585.1" /db_xref="GI:15155751"
CDS	3194 a 3242 g 2378 t	
BASE COUNT	2021 a	
ORIGIN		

	Query Match	52.2%	Score 45.4;	DB 1:	Length 10835;	
	Best Local Similarity	73.4%	Pred. No. 0.0044;			
Matches	58:	Conservative	0:	Mismatches	21:	Indels
					0:	Gaps
						0:
QY	8	tcttcctcgtcgtcgtcgcataccgacalcctcgcgagtgcgatcgtcgcaaacagcacg	67			
Db	6552	TCTTCTGTCGTGGTGCGCTGTTCGCCGCACATCGCGCGGTTTCACACCACAAGCAT	6611			
OY	68	tcggatccagaatgccgcgc	86			
Db	6612	TCCGTTTCGAAGCGCCCCC	6630			
RESULT	5					
LOCUS	AEO06000/c					
DEFINITION	Caulobacter crescentus section 326 of 359 of the complete genome.					
VERSION	AEO06000 AEO05673					
KEYWORDS	AEO06000.1 GI:13425102					
SOURCE	Caulobacter crescentus.					
ORGANISM	Caulobacter crescentus Bacteria; Proteobacteria; alpha subdivision; Caulobacter group; Caulobacter					
REFERENCE	(bases 1 to 12299)					
AUTHORS	Nierman,W.C., Feldblum,T.V., Laub,M.T., Paulsen,I.T., Nelson,K.E., Eisen,J., Heidelberg,J.F., Alley,M.R.K., Ohta,N., Maddock,U.R., Potocka,I., Nelson,W.C., Newton,A., Stephens,C., Phadke,N.D., Ely,B., Ely,B., Deboy,R.T., Dodson,R.J., Durkin,A.S., Gwinn,M.L., Hait,D.H., Kolonay,J.F., Smit,J., Craven,M., Khouli,H., Shetty,J., Berry,K., Uterback,T., Tran,K., Wolf,A., Vamathevan,J., Emoleva,M., White,O., Salzberg,S.L., Venter,J.C., Shapiro,L. and Fraser,C.M. Complete genome sequence of Caulobacter crescentus Proc. Natl. Acad. Sci. U.S.A. 98 (??), 4136-4141 (2001) 2 (bases 1 to 12299)					
JOURNAL	Nierman,W.C., Feldblum,T.V., Paulsen,I.T., Nelson,K.E., Eisen,J., Heidelberg,J.F., Alley,M.R.K., Ohta,N., Maddock,U.R., Potocka,I., Nelson,W.C., Newton,A., Stephens,C., Phadke,N.D., Ely,B., Laub,M.T., Deboy,R.T., Dodson,R.J., Durkin,A.S., Gwinn,M.L., Hait,D.H., Kolonay,J.F., Smit,J., Craven,M., Khouli,H., Shetty,J., Berry,K., Uterback,T., Tran,K., Wolf,A., Vamathevan,J., Emoleva,M., White,O., Salzberg,S.L., Shapiro,L., Venter,J.C. and Fraser,C.M. Direct Submission Submitted (31-JAN-2001) The Institute for Genomic Research, 9712 Medical Center Dr, Rockville, MD 20850, USA					
MEDLINE						
REFERENCE						
AUTHORS						
FEATURES	source					
	1..12299	/organism="Caulobacter crescentus"				
	/db_xref="taxon:69394"					
	102..1913	/gene="CC3393"				
	102..1913	/gene="CC3393"				
	/note="Identified by match to PFAM protein family HMM PF00441"					
	/codon_start=1					
	/transl_table=1					
	/product="acyl-CoA dehydrogenase family protein"					
	/protein_id="AAK5355.1"					
	/db_xref="GI:13425103"					
	/translation="MATUGAPLKGARPGSPWTFRAPVRDLASLRHVAGFERLADA FPEADADVAVALEAGAFASDIAPILNRQDILVGARLENSVPRAPGFADAYQOFA GGMTSLAAPRGGOGLPKTLEVAVLEWQAANNAFGICPMISGALFALBGSPDG KAKIKLVSGEMTGMNLTPPOAGSIDLAATTMPDPGDGMRITGQKIYTWDHHT AADNVHLVARLTDPADPAGVGISFLPAKVLDVDGKLGPANMLRASLSKHGIHG SPNCVLFEGAKARTNGISGLGIPNMFMMAAALONGTOGSVAIAERYOOALAFSOI RAOGRSANTGAYPSRLDPHPRYRTVLUMKRHHIAPRSICTSTAYADLARAAPDA RTAKARELLPTLAKAWSTDIGWVASOVQIHGGMGFIETGCAOHTRARLAPVT EGTGAQAIDIVGRKIMIEGOAAVDLMDIDTRDTIDLKASDLVSGRIRTAALDAAPA					
GENE	gene					
CDS						

gene  
CDS

complement(1984..2466)  
/gene="CC3394"  
complement(1984..2466)  
/gene="CC3394"  
/note="identified by match to Pfam protein family HMM PF00578"

/codon\_start=1  
/transl\_table=11  
/product="Abpc/Tsa family protein"  
/protein\_id="AAK25356.1"  
/db\_xref="GI:13425104"  
/translation="MATKVGDTPLRATFMSTACGAPISDDIFKKKTVALFVPRPAFTPTPQADLHFEKERADELKAGVDSTIVCSYNDVVMKWMDGSDIDSEVLTIAGNGDFFTKAIGLDHPDGSKFGMGARSORYSLVAKDGVVTQLDHVEDAQFRVSAYELLEDL."

gene  
CDS

2595..3182  
/gene="CC3395"  
2595..3182  
/gene="CC3395"  
/note="identified by Glimmer2; putative"  
/codon\_start=1  
/transl\_table=11  
/product="conserved hypothetical protein"  
/protein\_id="AAK25357.1"  
/db\_xref="GI:13425105"  
/translation="MASLTDDDCGEFLIGRLLVAMPGETEDPERFYLYICAHDEDAAMGLAVRVEEGLTVFELNRLGVRSKEIOAPSIDLILSGPERERGVLHTHDPESSPOSALTVPVADGTALRTDALDMASAIKRPKSLLAGVGWPGGLEDELDRNVNMCADEGLTFDDHEHKMTALKLGKITADHLSATAGA"

gene  
CDS

complement(3250..4917)  
/gene="CC3396"  
complement(3250..4917)  
/gene="CC3396"  
/note="identified by match to protein family HMM"  
/codon\_start=1  
/transl\_table=11  
/product="GGDF family protein"  
/protein\_id="AAK25358.1"  
/db\_xref="GI:13425106"

gene  
CDS

"MSFTGRIRIMDATATLEALGAADVAlMEPETDLRIKLNGAARALGPDPPECSSAFRALIAPORQAQEYLKPREGSEVVANRVRGCTETLCMRVCNALTEEVRAAGVAEPFKPSASELCDITLDRREFLARERLAOESTHOJVYDDLBRILRNPLAEGERADIYLAALGSRLAAPPASQIIIGEGDEFVICOPLEGESDVLRSAELOPVAGCDIHPIPTSIGVASNEGGLDAPDAELLRRALEAVEAAAARGGAARVGRMETGSLTALAENDLRKAIGKETTPFPQPIVRISTALSFEALANIIPRGMIPREPELPLEEGLMSELGAHHMHAAQULSTRANHPRMGMLIVSNISTCEIDTRPCGVADVEATELVNRNLPRGALKLEVESDIWRDERPAVILLKTLRDLAGSIALLDETCTGGSSSYTLRFEDFLKIDRYFVRTMGNNAGSANIKVRSVVKLGODLDIEVAACEVENNAEMAAHQSCGYGGOGGFYAPALSPQEAELYINLAYVDGAAPVRKG"

gene  
CDS

complement(4952..5530)  
/gene="CC3397"  
complement(4952..5530)  
/gene="CC3397"  
/note="identified by match to protein family HMM"  
/codon\_start=1  
/transl\_table=11  
/product="ribosomal-protein-alanine acetyltransferase, putative"  
/protein\_id="AAK25359.1"  
/db\_xref="GI:13425107"

gene  
CDS

"MSVLIPLFSPPRLRLQGRAVYLRPQRORDYAPMAALSRKSRAETLEPMPTGWPHDLSRAAFRRLAAVARLELEGEAVPFIFRED DALIGALRLFHVRRGSLTGITWGEEPYBROGHMAADVELLIFAFHGJHRLEAACMPENHASAALLAKCGFSDEGNHYUKINKCAMPDHRLFGLVAB"

gene  
CDS

(5546..6819)  
/gene="CC3398"  
complement(5546..6819)  
/gene="CC3398"  
/note="identified by match to Pfam protein family HMM PF00675"  
/codon\_start=1



Matches	58; Conservative	0; Mismatches	24; Indels	0; Gaps	0;
QY	5	cgatctccgtcgtgtcgcgcgtacgcgcacatctgcgcgtgacgagatgctgcggaagaacgc	64		
Db	9641	CGCTTCTCTGATCGCTCGCCTGATCGCCCTCATACCGCGCCCTTCACGCCCCCAAGC	9582		
QY	65	acgtcggtttcgagatgacgcgc	86		
Db	9581	ACTTCGCGCTTCGAAGCGCCGC	9560		
RESULT	6				
LOCUS	AP003011				
DEFINITION	AP003011 346510 bp	DNA	ECT	15-MAY-2001	
ACCESSION	Mesorhizobium loti DNA, complete genome, section 18/21.				
VERSION	AP003011 BA000012				
KEYWORDS	AP003011.2 GI:14026664				
SOURCE					
ORGANISM	Mesorhizobium loti (strain:MAFF303099) DNA.				
REFERENCE	Mesorhizobium loti				
AUTHORS	Bacteria; Proteobacteria; alpha subdivision; Rhizobiaceae group; Phyllobacteriaceae; Mesorhizobium.				
TITLE	1 (sites)				
JOURNAL	Kaneko,T., Nakamura,Y., Sato,S., Asamizu,E., Kato,T., Sasamoto,S.,				
MEDLINE	Matnabe,A., Idesawa,K., Ishikawa,A., Kawashima,K., Kimura,T.,				
REFERENCE	Kishida,Y., Kiyokawa,C., Kohara,M., Matsumoto,M., Matsuno,A.,				
AUTHORS	Mochizuki,Y., Nakayama,S., Nakazaki,N., Shlimpo,S., Sugimoto,M.,				
TITLE	Takenuchi,C., Yamada,M. and Tabata,S.				
JOURNAL	Complete genome structure of the nitrogen-fixing symbiotic				
MEDLINE	Bacterium Mesorhizobium loti				
REFERENCE	DNA Res. 7 (6), 331-338 (2000)				
AUTHORS	2 (bases 1 to 346510)				
TITLE	Kaneko,T.				
JOURNAL	Direct Submission				
MEDLINE	Submitted (05-DEC-2000) Takakazu Kaneko, Kazusa DNA Research				
REFERENCE	Institute, The First Laboratory for Plant Gene Research; Yana				
AUTHORS	1532-3, Kisarazu, Chiba 292-0812, Japan				
TITLE	(E-mail: kaneko@kazusa.or.jp;				
JOURNAL	URL: http://www.kazusa.or.jp/rhizobase/				
MEDLINE	Tel:81-438-52-3935(ex.2338), Fax:81-438-52-3934)				
REFERENCE	On May 11, 2001 this sequence version replaced gi:11994986.				
FEATURES	Location/Qualifiers				
SOURCE	1. 346510				
gene	/organism="Mesorhizobium loti"				
CDS	/strain="MAFF303099"				
gene	/db_xref="taxon:381"				
CDS	/complement(242..1264)				
gene	/gene="ml17087"				
CDS	/complement(242..1264)				
gene	/gene="ml17087"				
CDS	/note="similar to lipopolysaccharide core biosynthesis				
gene	glycosyl transferase Upse"				
CDS	/codon_start=1				
gene	/transl_table=1				
CDS	/protein_id="BAB53261.1"				
gene	/db_xref="GI:14026665"				
CDS	/translation="MAGALACAAFLVSDVHLPRITLADRPGLADIRKTONLRVLR				
gene	REIYRFQHDVWESDVNADIVLIFWQDILKPLRESEVLARSRLVIGTCSHEL				
CDS	EEFEPFGATLHLRLARAVFYNNRKLEHYAPLRLIPVHYTPNGVDITFFCRPEFOP				
gene	YSGELRVKAGSLGNHGFARHGFHDVTEFAVAANPGARLLTLIREGHMNHHEMDLFY				
CDS	RLLDVIVVCAKSEGEFNPCLAAAGCIPLVITRVGMPELVODGDLFFPDGTAEGLA				
gene	DLTALIRDTPLPSLRSRMAARMIETIRGMDRVOAENYAHFTSLTLAGVAEVEVRRAAS				
CDS	DVPGGY"				
gene	complement(1228..2265)				
CDS	/gene="ml17088"				
gene	complement(1228..2265)				
CDS	/gene="ml17088"				
gene	/note="similar to lipopolysaccharide core biosynthesis				
CDS	glycosyl transferase Lpsc"				
gene	/codon_start=1				
CDS	/transl_table=1				
gene	/protein_id="BAB53262.1"				

gene  
CDS  
/db\_xref="GI:14026666"  
/translation="MALIAGEHIDLGGVSTKCGALAMPSSLIRLVI SAKFSNCALL  
ERPAHSPAVNEGLIRVELNLTTHAARENSNRVAVALTTNREERFVAGCLENTFH  
OGVOYUJLIDNOSTDRTLEIAQRYAGAGLIGFESIPHNGWCWQLRKKEBLEPQSEA  
DMHMLADDETHLPTPTSSLSIYAADAALGFNAISSEFTIPTREAPDHNPQYQ  
HTRIRYVPRFSPCHVRAVKKQDCPTEIAMSCHRYRFGPRLSEPRMKHYETL  
SAEHAKRYGGRRLSEVYDGRGHGMRPLRPEDIDLPITQVRTTATDDIDLEANP  
WLHMLHAKYAS"  
complement(2648, .3787)  
/gene="ml17089"  
complement(2648, .3787)  
/gene="ml17089"  
/note="similar to mandelate racemase"  
/codon\_start=1  
/transl\_table=1  
/protein\_id="BAB53263.1"  
/db\_xref="GI:14026667"  
/translation="MSLSHFIRITRFOPARDVIGDSQVADVYVAALLESBSGEVG  
LGEITLNPDPDOEIESVEHEVMPISKONRAIALVHRVNRPGNGQAYSLPHE  
AFVQVAMPLAKREGGLRPLHYLIGSRNRKVRVYAGGLDPAFVLSFHAISIGS  
AKIYVGRHDPDRDLRLLELKTGCVPSKMDIPNPAWMSKEKLTVAIRREGHLS  
LWVEDPILRHHDGIRTLRRAVVTQINSQYIDLOSKRLLEHNAHDILNNGCVYD  
VYMRIGLAEGISIPISIGNTLEAGVHAVALPEAEENLVSFQNFDLVQPIEIRDG  
YVAYDRPHGLVYSEKARGEMSRRLRLSELGAADENPLPAK"  
3870, .4556  
/gene="ml17090"  
3870, .4556  
/gene="ml17090"  
/codon\_start=1  
/transl\_table=1  
/product="transcriptional regulator"  
/protein\_id="BAB53264.1"  
/db\_xref="GI:14026668"  
/translation="MAGKAKTDSSPLVAARMSDIAYERILEGDFDRVPAGATISON  
ELSAIVDPVAPRIADLRLEAEGLITHPRSQIOPAPRGELIRATYOPRSYIERAA  
VYRFDEQGEAVMALEIRHSHRLRVRDOLCKQLEPMALGELGAVITALRNP  
LIDSTYRRHNTLRLLDRKITTPTWVTLKEHDLILEAGTRNADAERKLAHQ  
AMORSRLGN"  
4684, .5922  
/gene="ml17091"  
4684, .5922  
/gene="ml17091"  
/codon\_start=1  
/transl\_table=1  
/product="sugar ABC transporter, periplasmic binding  
protein"  
/protein\_id="BAB53265.1"  
/db\_xref="GI:14026669"  
/translation="MOLNKRQFIQMSALAMAGHTRGARASATISYWHHTSOSEMAS  
LIKITIGEPANADYVYTOENIPNEYAKYKSYAVLAGREDPTAMVLAERADITLAME  
GLIDILDRNGKRGKANFENRNTGTLSNGAIVAPARATYMDITRADYIEEGLGAG  
PRNPEDEISACRLTDPKSGRYAFMGAGGIVFYVDVAEGSPYIKQGPDAIDR  
AAALEITFYSDPLFEKRYVPSAPNDYSROIEMEGFRGTQAMVWHHTSINTLSAM  
KQKQESTPSPGCKRAHARVAYAGNIMKEDITEAMRWVSFGQTDALITLLESTG  
GEPFSAALADLPRTGNPTIYHAAVETLEFGRLPNNFVGAAGMSBNVNPAPFAVLITG  
OSTPBRADNPTIGLEALNR"  
5979, .6932  
/gene="ml17093"  
5979, .6932  
/gene="ml17093"  
/codon\_start=1  
/transl\_table=1  
/product="ABC transporter, permease protein"  
/protein\_id="BAB53266.1"  
/db\_xref="GI:14026670"  
/translation="MLSHRGSMTPSSSGYRPLRALGELSETRYMYWLLILSLPWIIL  
VIAPPLYMALISFREMRLTRPGLGWYGFKYAAMSDPWFALAKTVMVVAAVV  
LEALGILAAPTLNRDLPGTKLEAVLILPLVPLVWGHMALLDLRLGVINDVLA  
RGCVTQIRAFADPRTALAAATVIVAAHMGPEFALLILLAGKSTIPEDLYAAAVDQ  
GVSQGLRTLTPMLKTYIAAATVILRYIGVYNSPDLLITLGGPDAIYOVLSTVAFQ  
AKNEFEGTASLSVMEVYILMIFATVIRLSRITKGDDP"  
6925, .7770  
/gene="ml17094"

Query Match	Best Local Similarity	50.1%: Score 43.6; DB 1; Length 346510;
5	cgatcttcctgctgtcgtcgtcatccgcattccgcggtgctgcatgtccgcgcaaaacg	64
58:	Conservative 0; Mismatches 24; Indels 0; Gaps 0;	
gene	6925. .7770 /gene="mlr7094" /codon_start=1 /transl_table=11 /product="ABC transporter, permease protein" /protein_id="BAB53267.1" /db_xref="GI:14026671" /translation="MIESKRRRRVIANIAFYFLATLSMFCISGPMIMWFELSLKYBAD IYNQMOYIPRTPFNYYAIFQSGEPRLIANSIVTFETVIAVATGTLAAVARS FAFRGTQIMGLIYIRMPAYMMIPIUYVMVGLVDSREGALVAFSLIPLFW MKGEFDSAPKELESARIDGSTRGAMVRIYMPLRNGLVASSYVIAIAANREFLA LMTITQGGSTRMPVGLQLMVEFOLPVGVLAAIGLISLIFVLVFAIVORTMYOGLTA GAIKG"	
CDS	7781. .8860 /gene="mlr7095" 7781. .8860 /gene="mlr7095" /codon_start=1 /transl_table=11 /product="sugar ABC transporter, ATP-binding protein" /protein_id="BAB53268.1" /db_xref="GI:14026672" /translation="MATSLIKDVRSKGSVQVLAHGVDIIELEDGSPVLILGSGCGKST LNIITAGLDDTSQCDLILIGRSVYDLAPKRNITAMPOSAIATPMSVENEIGFLM RGVDAQQRKRAVDDARLQILSHLDNRPNLSGQQRORAMARATYRDEPLFLPEP LSLMDAKLRVEMRETEIKRLHERLGTSLVYTHQILEMTLRAVAAKGGHIOQFADP RTIYERPANVFAVSFTGSPAMNPLPGRLVSDAGVPRPAASGVSAIQAGKRTVDSPE VIGIERPEELKAVAGRDEASVTGVIVDEPTNPVVOGDOSI IARLPBPAGRRN EPFLIADPASVNLDPGSEHRIDI"	
gene	8864. .9082 /gene="msr7097" 8864. .9082 /gene="msr7097" /note="unknown protein" /codon_start=1 /transl_table=11 /protein_id="BAB53269.1" /db_xref="GI:14026673" /translation="MDAKALLAORGFCHRPMSRKRRCRRDRRHRLQRLMAVAEAY AGGRIVGEINQAMIPMWRIRIARQART"	
CDS	9079. .10692 /gene="mlr7098" 9079. .10692 /gene="mlr7098" /codon_start=1 /transl_table=11 /product="GMC-type oxidoreductase" /protein_id="BAB53270.1" /db_xref="GI:14026674" /translation="MSETSPFGDYDYITVGAGAGCVIANRLSADPPNNRLILEAGSS DRYVWDIPITGILFCMGNPRTDMMKTEPEAGLGRSLANPRGKVLGCGSVNMATY RGQAAIDDMRQAGNMGMSDDLYLPKRSFDYHGGSAHMGSGEVRERQISMPFI LDARDAEELGITPTGDDFNITGTREGSGYEYVORLRYCNNTSKAFLRGLIAARNLAV VTGAEIRLEFGDGRKRVGGLLRQGDRLRHRSAGOTIILAAGINSPKILLESGRPLD LSLSGVVNRHALPGVGENLQDDHQLRTVEVYANATINERHNILISAAGAVEALRK SGPLSMASPOLGIFARSDDLATPDLVEYHVOPLSTDRLGEPLHREFPVTYVSCVLRPE GRCVHMASSDPREATKIRPNYISTADRSVAASLRHVSIMKARAVARFAPEMDLP GTQVDSDDILIRKAGDIGTTFEHPGTCCKMGSDMAVVDRLRVHGLAGLRVVDASIM PTIYSGNTNSPVYIAEKAEMILLDRR"	
gene	10929. .11801 /gene="mlr7099" 10929. .11801 /gene="mlr7099" /codon_start=1 /transl_table=11 /product="esterase" /protein_id="BAB53271.1"	
CDS		





CDS		/gene="orf157" complement(7350..7823) /gene="orf157" /note="ORF157" /codon_start=1 /transl_table=11 /evidence=not_experimental /product="unknown" /protein_id="AAF78797.1" /db_xref="GI:8708906" /translation="MTSASVYPLLAIGASHHYLYAAHRDADACARLHGDPPTAAH IAEARAECAYVLETRHTSASLTIOENADPSVLDTTLASLAEANGWHDTEGPP DMPEGHKTMILQTSLOVPVNLGRALGTWQAIIYLHRAHPRRREVLIQTGAN"
gene		/gene="Orf235" complement(7853..8560) /complement(7853..8560) /gene="Orf235" complement(7853..8560) /gene="orf235" /note="ORF235" /codon_start=1 /transl_table=11 /evidence=not_experimental /product="unknown" /protein_id="AAF78798.1" /db_xref="GI:8708907"
CDS		/translation="MPCGRADYPARRRGCEDETSVVYHNIMAFSGAGNGDGQR SKCARIRARRKRRLQHIDADPRGAAGSTPAPARDLHOERRRGLLGARACARAR KDRPADRSRAEGTAGRDGVCKIDEALPKSPDKADPKSPATODRPAERPE RRSDVDYVGKGLASADRGGPRABEGRWGEDLSLSWRKHIRGLISRSALPPEFAH VTIRI RGNSSWLPCAIAH" 8893..9240 /gene="uncf" /note="orf115" 8893..11493 /gene="Fo ATPase operon" /note="constitutively expressed" 8893..9240 /gene="uncf" /note="ORF115: initial gene of the Fo ATPase operon" /codon_start=1 /transl_table=11 /evidence=not_experimental /product="unknown" /protein_id="AAF78799.1" /db_xref="GI:8708908"
gene		/translation="METAIIDRPNRVRFPNNGSELIISCPNFRAATSRRPKAGDDGR AARSAMALCFRLSELVAGVVGAGIGWGDFDLISTSPGFIVFLILGVAGVNVV
gene		
CDS		
Query Match	48.3%; Score 42; DB 1; Length 31495;	
Best Local Similarity	69.5%; Pred. No. 0.033;	
Matches 57; Conservative 0; Mismatches 25; Indels 0; Gaps 0;		
OY 5	cgaattcttcgtcgttcgcatccgcacatccctcgcgagtgcatgtcgcgcaaacgc 64       Ccatttctctctggctgcacctgttcctgcctgatgccggccacttcacggcgaagcagc 20295 	
Dd 20354	aatctcgattccagatgagcgcg 86 	
OY 65	aatctcgattccagatgagcgcg 86 	
Dd 20294	atctcgacctcgagttcgccgc 20273	
RESULT 9		
AE004449		
LOCUS	AE004449 12829 bp DNA BCT 30-AUG-2000	
DEFINITION	Pseudomonas aeruginosa PA01, section 10 of 529 of the complete genome.	
ACCESSION	AE004449 AE004091	
VERSION	AE004449.1 GI:9945928	
KEYWORDS		
SOURCE	Pseudomonas aeruginosa.	
ORGANISM	Bacteria; Proteobacteria; gamma subdivision; Pseudomonadaceae; Pseudomonas.	
REFERENCE	1 /bases 1 to 12829)	

**AUTHORS** Stover, C.K., Pham, X.Q., Erwin, A.L., Mizoguchi, S.D., Warren, P., Hickey, M.J., Brickman, F.S., Hurnagle, W.O., Kowalik, D.J., Lagrou, M., Gardner, R.L., Goltzy, L., Tolentino, E., Westbrook-Wadman, S., Yuan, Y., Brody, L.L., Coulter, S.N., Folger, K.R., Kas, A., Lardig, K., Lim, R., Smith, K., Spencer, D., Wong, G.K., Wu, Z., and Paulsen, I.T.  
Complete genome sequence of *Pseudomonas aeruginosa* PA01, an opportunistic pathogen  
Nature 406 (6799), 959-964 (2000)

**JOURNAL** 20437337

**MEDLINE** 2 (bases 1 to 12829)

**REFERENCE** Stover, C.K., Pham, X.Q., Erwin, A.L., Mizoguchi, S.D., Warren, P., Hickey, M.J., Brickman, F.S., Hurnagle, W.O., Kowalik, D.J., Lagrou, M., Gardner, R.L., Goltzy, L., Tolentino, E., Westbrook-Wadman, S., Yuan, Y., Brody, L.L., Coulter, S.N., Folger, K.R., Kas, A., Lardig, K., Lim, R., Smith, K., Spencer, D., Wong, G.K., Wu, Z., Paulsen, I.T., Reizer, J., Sater, M.H., Hancock, R.E.W., Lory, S. and Olson, M.V.  
Direct Submission  
Submitted (16-MAY-2000) Department of Medicine and Genetics, University of Washington Genome Center, University of Washington, Box 352145, Seattle, WA 98195, USA

**TITLE** Location/Qualifiers

**FEATURES** 1..12829

**source** /organism="Pseudomonas aeruginosa"  
/strain="PA01"  
/db\_xref="taxon:287"  
219..947  
/gene="PA0102"  
219..947  
/gene="PA0102"  
/codon\_start=1  
/transl\_table=11  
/product="probable carbonic anhydrase"  
/protein\_id="AAG03492.1"  
/db\_xref="GI:9945929"  
/translation="MPDRRGVKSQSDPEESADALKRIVDFQHFREVEPEQALF  
KLLANSORPAMFICADSRIVPELTQSSPDLEFTRNGVNPVPGOMGVSSTAI  
EYAVIALGVHHTIYVNCDCGMAVILDPQLEPRPVPKMLRHAEEVARTVAVNDC  
GASHPTLVTFEENVVADLHFRPVSASRLASQQLIHGMVYDIESAQIRAVDAKQ  
GRPLDSEHPVPKATPAPRYLSS"  
1158..2729  
/gene="PA0103"  
1158..2729  
/gene="PA0103"  
/codon\_start=1  
/transl\_table=11  
/product="probable sulfate transporter"  
/protein\_id="AAG03493.1"  
/db\_xref="GI:9945930"  
/translation="MHVSKVPPSLRDTLPRLMASVVEIVLPLCMGIAIASGMPA  
KGLTGVIGAVIFLAGSPLOVSGPAGLAVIVELRTYGAAMLGPIILLAGAIOL  
LAGRLRLGCMFRTSPRAYVYCMILACILYLSOAHWMLDAPKASGIDILLAPFOAA  
FAALSLGMDSGDALGLGTITAYMGMDKIRPRLRPLRALGVSLLATLASMLA  
LDVRREVANLGEAIIVMLRPADLLADPSILLAAVVAFTASAEILLSAADVRLH  
DPRSDMRELSAGCVGNMLCGLALPMTGVIVRSANVNGAETLLSAFHLIML  
AFVLLGSLRLIIPVASLAGVLYTVKLVDFKALGNSNVNAGATRYAATLAIIV  
TLLTGVIGVAFPLTLIKVILKARLKILARTGEHEALRLSGAATFLKVALSVLD  
EYKPTTLHVPMDNISYDHAEMELLEDMGRNAPQGSLSVIERALKRRLGRLRGS  
VGLGARNGAASPC"  
2866..3462  
/gene="PA0104"  
2866..3462  
/gene="PA0104"  
/codon\_start=1  
/transl\_table=11  
/product="hypothetical protein"  
/protein\_id="AAG03494.1"  
/db\_xref="GI:9945931"  
/translation="WVDFVQPVGEGLADQLVGHQVYGVQVQDCLAEGVAVVHF  
FHQAVRQVGEVGVAAVVAAGORRGEELGDIQHLRDLPLDAGFAGHAGVGEV  
RILFDVGEAGEVYRVYALRLDQAMQVQAGAOISQVTAAGFAEVRVLLDEPTEGS  
GVHGVSHTTGMVKKRRNITAGARRLRDRDHNRFVYVNG"  
3726..4850

**gene** /gene="coxB"  
/note="PA0105"  
3726..4850  
/gene="coxB"  
/codon\_start=1  
/transl\_table=11  
/product="cytochrome c oxidase, subunit II"  
/protein\_id="AAG03495.1"  
/db\_xref="GI:9945932"  
/translation="MLRHRVMMGPIILLISAQANAAMTVNMAPGATEVRSVFDLHM  
TPIWCVIGVLFEGAMFMSLVYHRSTQGRPHHRETTVEILMTPVPEVLLYVMAV  
PATRLILHIVDSEPLDVQVGYQWQWYKYGODVEYTESLAPPODINRQAKDE  
HYLEADDEPVLVPGTKVRFLLTSSDVHSMVVPFAVRDAIPGVNDAMRWDEPG  
IYRQCAELCGKHGCHMPYLVVYVKPFAEEDQMLAKKEBAKVELTSEKTEBLVA  
RGQVYHTTCAACHQAGGMPMPALGSKIVTGPKEHLLEVVFNGVPTAMAFG  
KQNEVDIAAVITYERNANGNDGQVTFKDVAAKQKQ"  
4860..6452  
/gene="coxA"  
/note="PA0106"  
4860..6452  
/gene="coxA"  
/codon\_start=1  
/transl\_table=11  
/product="cytochrome c oxidase, subunit I"  
/protein\_id="AAG03496.1"  
/db\_xref="GI:9945933"  
/translation="MSAVITDTPDHAGDHHGHPAKGLMRVLTNNHDKIGTLVLMFSF  
MPLFGSSNMAVIRAEIPOPGLQIVPAFFNQMTWGLINVGAVNPAFVGLANMI  
PLMTGAPDALPFRNNPFRVLLPAAAGLIVSTLFMRGCGPNCMPYAPLSTFAPHS  
VTFEIPALHLAGISSIMGATINVTATLLNRAPOMTLMKRLPLVWTLITAFLLIAMP  
VLAVVYTMGLMDIHFGTSEFSAAGGDPVLFQHVFEFHPREYIMILPAFASVSLI  
PFPARKPLFEGYTSNMYATASIAFLSEVVAHNFVVGIVTGLFPMVATMLIAVPTG  
KVFNWYTTMECSLTFEETPMLEFAVAVILFTTGSGFSLMLIAPADFOYHDTYEVVA  
HFVYLVPGAIFGIPASAYYMLPKMTGHWIDETGLHFMMSICGNLAEFFMHFVGL  
AGPRRIPDPYNLOPADPNMYSIGAFMFGTQLLFLIYIKCRGSKPAPAPMWDAG  
GLEMSLSPAPVHTFSTPREVK"  
6463..7017  
/gene="PA0107"  
6463..7017  
/gene="PA0107"  
/codon\_start=1  
/transl\_table=11  
/product="conserved hypothetical protein"  
/protein\_id="AAG03497.1"  
/db\_xref="GI:9945934"  
/translation="MSDAKYDTRRLVGRLLVTVLMFAFGFALVPLDYWCALGTING  
KTGSAVSGEOVDVGRVKKVQFMTSNNDIMWEPFSAGDOLVHFGAVNOMVYARN  
PSDKPRMRAQIPSTAPAEAAVYFKTECFQTOQVLOPGBSIMPVRFYDRDLPDV  
RHVLTATLTFDITARKPPVPAQR"  
7028..7915  
/gene="coiII"  
/note="PA0108"  
7028..7915  
/gene="coiII"  
/codon\_start=1  
/transl\_table=11  
/product="cytochrome c oxidase, subunit III"  
/protein\_id="AAG03498.1"  
/db\_xref="GI:9945935"  
/translation="MASHENYVPAQSKWPIIASIGLVTVFGLGTWPNDLNAGHRES  
HGPIVTFVGGIITAYMLFGWGNVIRSRAGLISAOQDRFSRGMSEFTSEVFEAA  
FEGALFVVRHFGAGPLGEGKAKGVANHLMNRFQVSRPLDTPPDKLIPPSAVIEPFW  
LPLINTILVTSSTVTFHAIHALKKNRGRPLKMLALVLLGTAFLILOEEVYHYN  
ELGLTGAGIYGTFFMLTGFHGAHVTLGALLIGLIMLIRLRGHFDEHHPGEASW  
YMHFVDVWIGLFIYVYV"  
complement(7931..8140)  
/gene="PA0109"  
complement(7931..8140)  
/gene="PA0109"  
/codon\_start=1  
/transl\_table=11  
/product="hypothetical protein"  
/protein\_id="AAG03499.1"

TITLE	JOURNAL	PUBMED	REFERENCE	AUTHORS	COMMENT
Renard, C., Thebaud, P., Vandenbol, M., Weidner, S. and Galibert, F. From the Cover: Analysis of the chromosome sequence of the legume symbiont <i>Sinorhizobium meliloti</i> strain 1021 Proceedings of the National Academy of Sciences of the United States of America. 98 (17), 9877-9882 (2001)		11481430			
Gouzy, J.		2 (bases 1 to 294800)			
Direct Submission Submitted (26-JUL-2001) Gouzy J., Submitted on behalf of the MELLIO EU Consortium					
MELLIO EU Consortium:					
Laboratoire de Biologie Moléculaire des Relations Plantes-Microorganismes, UMR215-CNRS-INRA, BP27, F-31326 Castanet, France, Laboratoire de Génétique et Développement UMR6061-CNRS, Faculté de Médecine, 2 avenue du Pr. Léon Bernard, F-35043 Rennes, France, GARC GMBH, Fritz-Arnold-Str. 23, D-78467 Konstanz, Germany, Universitätsklinik Bielefeld, Biologie IV (Genetik) Universitätsstr. 25, D-33615 Bielefeld, Germany, Unité de Biochimie physiologique, Université Catholique de Louvain, Place Croix du Sud 2, Bte 20, B-1346 Louvain-la-Neuve, Belgium, Unité de Microbiologie, Faculté des Sciences Agronomiques de Gembloux, Avenue Macheval 1u1n 6, B-5030 Gembloux, Belgium. E-mail: Jerome.gouzy@toulouse.inra.fr http://sequence.toulouse.inra.fr/meliloti.html.					
Location/Qualifiers					
1. 294800					
/organism="Sinorhizobium meliloti"					
/strain="1021"					
/db_xref="taxon:382"					
/complement(147. 1382)					
/gene="SMC04259"					
complement(147. 1382)					
/gene="SMC04259"					
/function="cell processes; transport of small molecules"					
/note="Product confidence : hypothetical					
Gene name confidence : hypothetical					
predicted by Codon usage					
predicted by Homology					
predicted by Framed"					
/codon_start=1					
/transl_table=11					
/evidence=not_experimental					
/product="PUTATIVE PERIPLASMIC BINDING ABC TRANSPORTER					
SIGNAL PEPTIDE PROTEIN"					
/protein_id="CAC46508.1"					
/db_xref="GI:15074951"					
/translation="MNLRLFLAALGAPALPEGAASATDLETVHHMTSGEAAVAEL					
AKAFDITGKMKWDGACJAGSGTARPMIMISITITGGDPMAQTQFHNGRAEELVQAGLMR					
DLTIDITRKNKTIKVPSSLLDSCTTEGKTYICAPVNIHSMQWMLNSAAKRAQGEVPE					
KNNDEVAAPALEKAGIVPLAVGGPQWNGAFVLMVAIAKRENEEYFAQKDEVP					
AAGPELTACPLGAVNEVISTGDAFEPFLIEDEKRAQEVLAATSLTKPTQVAFNLK					
KGLSPRGVNDLAANDCMKRGDIIAKGVIGQTDQLLSADSOQKQKDELFSPFANH					
SMPEAQKREFNDITIAAD"					
1858. 2928					
/gene="SMC04260"					
1858. 2928					
/gene="SMC04260"					
/function="miscellaneous; not classified regulator"					
/note="Product confidence : putative					
Gene name confidence : hypothetical					
predicted by Codon usage					
predicted by Homology					
predicted by Framed"					
/codon_start=1					
/transl_table=11					
/evidence=not_experimental					
/product="PUTATIVE TRANSCRIPTION REGULATOR PROTEIN"					
/protein_id="CAC46509.1"					
/db_xref="GI:15074952"					
/translation="MTGMDGKTRNRTATPAVEGGRPILKTAFTMGIVTVYRAK					
DAPDIGAETKREVRIVAKQIGQPNRAGVRLRTGKTNVLSIVLTLEEIMGITSPIV					
GTEIILAQOYHVLVTPYSSADPLGPRIIYIDTGAADVIIISRTPEPNDPRVILTER					







TITLE		Direct Submission	
JOURNAL		Submitted (10-JAN-2001) Genomics, Monsanto, 800 North Lindbergh Blvd., Creve Coeur, MO 63167, USA	
COMMENT		Derived from rice genomic sequences generated from the Monsanto Rice Genome Sequencing project. Please see <a href="http://www.rice-research.org">http://www.rice-research.org</a> for more information. The sequence data were produced primarily in the laboratories of Dr. Leroy Hood at the University of Washington in Seattle.	
FEATURES		Location/Qualifiers	
SOURCE		1..227	
repeat_region		/db_xref="taxon:4530"	
		1..227	
		/note="microsatellite MRC5780"	
		/rpt_type=tandem	
BASE COUNT		23 a 89 c 75 g 40 t	
ORIGIN			
Query Match	34.9% Score 30.4; DB 8; Length 227;		
Best Local Similarity 61.2%;	Pred. No.1.3e+02;		
Matches 49;	Conservative 0; Mismatches 31; Indels 0; Gaps 0;		
OY 6	gattctctgtctgtgtcctgcatgcagatccctgcgcggtgcatgtcgcacaacagca 65		
Db 93	GCCTCATCCCTCCTCCTCCTCCTCCTCCTCCTCCTCCTCCTCCTCCTCGCGCCTTGTCGTCGACAGACGACGA 152		
OY 66	cgtgcggttcgagatgcgcg 85		
Db 153	CGTCGTGCTCGAGCTGGCG 172		
RESULT 12			
LOCUS	MACCYO3A 714 bp mRNA PRI 15-AUG-2001		
DEFINITION	Macaca mulatta cytochrome oxidase III mRNA, 3' end.		
ACCESSION	M85148		
VERSION	M85148.1 GI:342260		
KEYWORDS	cytochrome oxidase III.		
SOURCE	Macaca mulatta male adult cerebral cortex cDNA to mRNA.		
ORGANISM	Macaca mulatta		
	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Cercopithecoidea; Macaca.		
	1 (bases 1 to 714)		
REFERENCE	Chandrasekaran,K., Stoll,J., Giordano,T., Attack,J.R., Matocha,M.F., Brady,D.R. and Rapoport,S.I. Differential expression of cytochrome oxidase (COX) genes in different regions of monkey brain J. Neurosci. Res. 32 (3), 415-423 (1992)		
TITLE			
JOURNAL			
MEDLINE			
PUBMED	93059501		
FEATURES	1279190		
SOURCE	Location/Qualifiers		
	1..714		
	/organism="Macaca mulatta"		
	/db_xref="taxon:9544"		
	/sex="male"		
	/tissue_type="cerebral cortex"		
	/dev_stage="adult"		
	1..714		
	/gene="COIII"		
	<1..236		
	/gene="COIII"		
	/note="longest ORF"		
	/codon_start=3		
	/product="cytochrome oxidase subunit III"		
	/protein_id="AA802131.1"		
	/translation="SSYSLNNIMLGVISILLLHHPNTRPTNYIMPMTRHCTR KHLPRPHNTRPKMLRGITLVEAFVEGF"		
gene			
CDS			
	200 a 224 c 97 g 193 t		
BASE COUNT			
ORIGIN			

```

Query Match Similarity      34.7%; Score 30.2; DB 9; Length 714;
Best Local Similarity      60.2%; Pred. No. 1.2e+02;
Matches 50; Conservative 0; Mismatches 33; Indels 0; Gaps 0;

QY      4      tgcattcttcgctgtcgtcgtcgtacatcgacatctgcgcgtgcatgtcgcgaacacag 63
           ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
DB      569      ttcacattctctcttctgtctgctccatccgacacactcactacacacattcagaccat 628
           ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||

QY      64      caccgtcgttcgagatgcgcgc 86
           ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
DB      629      CACTTCGGCTTCGAGACCCGCG 651

RESULT 13
LOCUS      AC019055      176383 bp      DNA      PRI      03-JUL-2001
DEFINITION Homo sapiens clone RP11-111P18, complete sequence.
ACCESSION      AC019055
VERSION      AC019055.5      GI:14589677
KEYWORDS      HTG.
SOURCE      human.
ORGANISM      Homo sapiens
               Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
               Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.

REFERENCE      1 (bases 1 to 176383)
AUTHORS      Waterston,R.H.
TITLE      The sequence of Homo sapiens clone
JOURNAL      Unpublished
REFERENCE      2 (bases 1 to 176383)
AUTHORS      Waterston,R.H.
TITLE      Direct Submission
JOURNAL      Submitted (30-DEC-1999) Genome Sequencing Center, Washington
               University School of Medicine, 4444 Forest Park Parkway, St. Louis,
               MO 63108, USA
REFERENCE      3 (bases 1 to 176383)
AUTHORS      Waterston,R.H.
TITLE      Direct Submission
JOURNAL      Submitted (09-AUG-2000) Genome Sequencing Center, Washington
               University School of Medicine, 4444 Forest Park Parkway, St. Louis,
               MO 63108, USA
REFERENCE      4 (bases 1 to 176383)
AUTHORS      Waterston,R.H.
TITLE      Direct Submission
JOURNAL      Submitted (10-SEP-2000) Genome Sequencing Center, Washington
               University School of Medicine, 4444 Forest Park Parkway, St. Louis,
               MO 63108, USA
REFERENCE      5 (bases 1 to 176383)
AUTHORS      Waterston,R.
TITLE      Direct Submission
JOURNAL      Submitted (30-SEP-2000) Department of Genetics, Washington
               University, 4444 Forest Park Avenue, St. Louis, Missouri 63108, USA
REFERENCE      6 (bases 1 to 176383)
AUTHORS      Waterston,R.H.
TITLE      Direct Submission
JOURNAL      Submitted (03-JUL-2001) Genome Sequencing Center, Washington
               University School of Medicine, 4444 Forest Park Parkway, St. Louis,
               MO 63108, USA
COMMENT      On Jul 3, 2001 this sequence version replaced gi:9755483.

----- Genome Center -----
Center: Washington University Genome Sequencing Center
Center code: WUGSC
Web site:http://genome.wustl.edu/gsc/index.shtml
----- Project Information -----
Center project name: H.NH0111P18
-----
Location/Qualifiers
1. 176383
/organism="Homo sapiens"
/db_xref="taxon:9606"
/clone="RP11-111P18"

```



51003 68891 contig of 17889 bp in length  
 68992 84339 contig of 15348 bp in length  
 84440 97577 contig of 11338 bp in length  
 97678 109261 contig of 11584 bp in length  
 109362 118150 contig of 8789 bp in length  
 118251 127677 contig of 9427 bp in length  
 127778 134661 contig of 6884 bp in length  
 134762 139706 contig of 4945 bp in length  
 139807 145253 contig of 5447 bp in length  
 145354 150450 contig of 5097 bp in length  
 150551 156033 contig of 5483 bp in length  
 156134 160013 contig of 3880 bp in length  
 160014 160113 contig of 3300 bp in length  
 160114 163443 contig of 3330 bp in length  
 163444 163543 contig of 100 bp in length  
 163544 166187 contig of 2644 bp in length  
 166188 166287 contig of 100 bp in length  
 166288 169598 contig of 3311 bp in length  
 169599 171442 contig of 1744 bp in length  
 171443 171542 contig of 100 bp in length  
 171543 173750 contig of 2208 bp in length  
 173751 173850 contig of 100 bp in length  
 173851 175411 contig of 1561 bp in length

1 29465: contig of 29465 bp in length  
 29466 29565: gap of 100 bp  
 29566 50902: contig of 21337 bp in length  
 50903 51002: gap of 100 bp  
 51003 68891: contig of 17889 bp in length  
 68892 68991: gap of 100 bp  
 68992 84339: contig of 15348 bp in length  
 84340 84439: gap of 100 bp  
 84440 97577: contig of 13138 bp in length  
 97578 97677: gap of 100 bp  
 97678 109261: contig of 11584 bp in length  
 109262 109361: gap of 100 bp  
 109362 118150: contig of 8789 bp in length  
 118151 118250: gap of 100 bp  
 118251 127677: contig of 9427 bp in length  
 127678 127777: gap of 100 bp  
 127778 134661: contig of 6884 bp in length  
 134662 134761: gap of 100 bp  
 134762 139706: contig of 4945 bp in length  
 139707 139806: gap of 100 bp  
 139807 145253: contig of 5447 bp in length  
 145254 145353: gap of 100 bp  
 145354 150450: contig of 5097 bp in length  
 150451 150550: gap of 100 bp  
 150551 156033: contig of 5483 bp in length  
 156034 156133: gap of 100 bp  
 156134 160013: contig of 3880 bp in length  
 160014 160113: gap of 3300 bp in length  
 160114 163443: contig of 3330 bp in length  
 163444 163543: gap of 100 bp in length  
 163544 166187: contig of 2644 bp in length  
 166188 166287: gap of 100 bp in length  
 166288 169598: contig of 3311 bp in length  
 169599 171442: gap of 1744 bp in length  
 171443 171542: gap of 100 bp in length  
 171543 173750: contig of 2208 bp in length  
 173751 173850: gap of 100 bp in length  
 173851 175411: contig of 1561 bp in length

NOTE: This is a 'working draft' sequence. It currently consists of 28 contigs. The true order of the pieces is not known and their order in this sequence record is arbitrary. Gaps between the contigs are represented as runs of N, but the exact sizes of the gaps are unknown. This record will be updated with the finished sequence, as soon as it is available and the accession number will be preserved.

## FEATURES

source	FEATURES
misc_feature	175412 175511: gap of 100 bp
misc_feature	175512 177277: contig of 1766 bp in length
misc_feature	177278 177377: gap of 100 bp in length
misc_feature	177378 178748: contig of 1371 bp in length
misc_feature	178749 178848: gap of 100 bp in length
misc_feature	178849 179978: contig of 1130 bp in length
misc_feature	179979 180078: gap of 100 bp in length
misc_feature	180079 181125: contig of 1047 bp in length
misc_feature	181126 181225: gap of 100 bp in length
misc_feature	181226 181791: contig of 566 bp in length
misc_feature	181792 181891: gap of 100 bp in length
misc_feature	181892 182903: contig of 1012 bp in length
misc_feature	182904 183003: gap of 100 bp in length
misc_feature	183004 184021: contig of 1018 bp in length
misc_feature	184022 184121: gap of 100 bp in length
misc_feature	184122 185147: contig of 1026 bp in length.
Location/Qualifiers	
1. .185147	
/organism="Homo sapiens"	
/db_xref="taxon:9606"	
/chromosome="18"	
/map="18q23"	
/clone="RP11-675M6"	
1. .29465	
/note="assembly-fragment"	
29566. .50902	
/note="assembly-fragment"	
51003. .68891	
/note="assembly-fragment clone_end:r7 vector_side:right"	
68992. .84339	
/note="assembly-fragment"	
84440. .97577	
/note="assembly-fragment"	
97678. .109261	
/note="assembly-fragment"	
109362. .118150	
/note="assembly-fragment"	
118251. .127677	
/note="assembly-fragment"	
127778. .134661	
/note="assembly-fragment"	
134762. .139706	
/note="assembly-fragment"	
139807. .145253	
/note="assembly-fragment"	
145354. .150450	
/note="assembly-fragment"	
150551. .156033	
/note="assembly-fragment"	
156134. .160013	
/note="assembly-fragment"	
160114. .163443	
/note="assembly-fragment"	
163544. .166187	
/note="assembly-fragment"	
166288. .169598	
/note="assembly-fragment"	
169699. .171442	
/note="assembly-fragment"	
171543. .173750	
/note="assembly-fragment"	
173851. .175411	
/note="assembly-fragment"	
175512. .177277	
/note="assembly-fragment"	
177378. .178748	
/note="assembly-fragment"	
178849. .179978	
/note="assembly-fragment"	
180079. .181125	
/note="assembly-fragment"	
181226. .181791	
/note="assembly-fragment clone_end:SP6 vector_side:right"	

Thu Jan 10 08:15:52 2002

us-09-712-768-7.rge

Page 18

```
misc_feature 181892..182903
              /note="assembly_fragment"
```

Query Match	Score	DB 2;	Length
34.38;	29.8;	DB 2;	185147;
Post Local Similarity	63.08;	Pred NO.	63;

Best Local Similarity 63.0%; Pred. No. 63;  
Matches 46; Conservative 0; Mismatches 27; Indels 0; Gaps 0;

Oy 1 gattcgatcttcctgctcgtccctgatccgcacacctcgcggagtgcagatgtcggaanaa 60  
||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||  
Db 70885 GATTCACATCATCAGAGTGTCTGCTCCCTCCGGCTCCAGAGGCATCCCATCTCTCAAAA 70944

```

Qy      61 cagcacgtcgtt 73
      || | | | |
Db 70945 CATAGCTTCTGCT 70957

```

Search completed: January 10, 2002, 00:57:09  
Job time: 3117 sec

[illegible]

Query Match	100.0%	Score 87	DB 22	Length 87
Best Local Similarity	100.0%	Pred. No. 3	7-20	
Matches 87	Conservative 0	Mismatches 0	Indels 0	Gaps 0
QY	1	gattcgatcttcctcgtctgctgctgatacgcacatccgcgcggtgagatgcgcgaana	60	
		gattcgatcttcctcgtctgctgctgatacgcacatccgcgcggtgagatgcgcgaana	60	
Db	1	gattcgatcttcctcgtctgctgctgatacgcacatccgcgcggtgagatgcgcgaana	60	

RESULT	2
AAH22961/c	
ID	AAH22961 standard; cDNA; 1837 BP

DT 17-SEP-2001 (first entry)

Human phosphatase (PP) encoding cDNA (clone ID 1723447CBI).

KW phosphatase; PP; Alzheimer's disease; Huntington's disease; human  
KW immune disorder; cancer; neotropic; immunomodulatory; cytosstatic;  
KW neuroprotective; anticonvulsant; ss.

OS Homo sapiens.

FH	Key	Location/Qualifiers
FT	CDS	1..1323

PN WO200153469-A2.

PD 26-JUL-2001.

PF 18-JAN-2001; 2001WO-US02088.

PR 21-JAN-2000; 2000US-01777719.

PR 25-FEB-2000; 2000US-0184959.

PR 17-MAR-2000; 2000US-0190142.

PA (INCY-) INCYTE GENOMICS INC.

[illegible]

RESULT	3
AAQ10190	
ID	AAQ10190 standard; DNA; 23666 BP

DT 27-MAR-1991 (first entry)

**Cephalosporin antibiotic biosynthetic genes.**

KW	cephalosporin; antibiotic;
KW	S-(L-alpha-aminoadipyl)-L-cysteine-D-; valine synthetase
KW	isopenicillin N synthetase; isopenicillin N epimerase;

KW deacetoxycephalosporin C hydroxylase; ss

OS Lysobacter lactamgenus.  
XX

ET	/lab
----	------

/ \* t a

FT	CDS	1780

FT	1070	1070	1070
FT	1070	1070	1070

/ \* t e  
/ 1 - b  
F T  
re

FT	cm <sup>-1</sup>	Label
1096		

/\*t=

FT	CDS	21186..22343
FT		/*tag= f
FT		//label= ORF 6
FT	CDS	4211..5443
FT		/*tag= g
FT	CDS	//label= ORF 7
FT		2887..4200
FT		/*tag= h
FT	CDS	//label= ORF 8
FT		6..2819
FT		/*tag= i
FT		//label= ORF 9
XX		
XX	PN	JF02291274-A.
XX		
XX	PD	03-DEC-1990.
XX		
XX	PF	10-JAN-1990; 90JP-0003762.
XX		
XX	PR	01-FEB-1989; 89JP-0024710.
XX	PR	10-JAN-1990; 90JP-0003762.
XX		
XX	PA	(TAKE ) TAKEDA CHEMICAL IND KK.
XX		
XX	DR	WPI: 1991-01854/03.
XX	DR	P-PSDB: AAR10145,RI0688-AAR10695.
XX		
XX	PT	Prepn. of cephalosporin series antibiotics - comprises culturing
XX	PT	transformant of microbe transformed by plasmid contg. new DNA
XX	PT	fragment
XX		
XX	PS	Claim 4; Fig 1; 67pp: Japanese.
XX		
XX	CC	A fragment of the sequence comprising at least one of the
XX	CC	cephalosporin biosynthetic enzymes listed in the KEYWORDS can be
XX	CC	cloned in a plasmid and used to transform microbes, such as
XX	CC	bacteria or yeast. Although DNA is preferably isolated from L.
XX	CC	bactangens, similar sequences could be obtained from other
XX	CC	bacteria containing genes coding for biosynthesis of cephalosporin
XX	CC	series antibiotics. See also AAQ10191-2.
XX		
SQ		Sequence 23666 BP; 3952 A; 8522 C; 7603 G; 3589 T; 0 other;
QY	Query Match	31.5%; Score 27.4; DB 12; Length 23666;
Db	Best Local Similarity	59.7%; Pred. No. 9.4;
Matches	46; Conservative	0; Mismatches 31; Indels 0; Gaps 0;
QY	9 ctctcgtgtgtctgtcctcgatccgcatctctgcgcggttcgcatgttcgccaacagcacgt 68	
Db	3837 ctctcgtgtgtgtctgtcctcgttcgaacacatcgcctcgtctgtcgtgcgaagtcatcgtgcg 3896	
QY	69 cgattcgaagatgcgcg 85	
Db	3897 cgcttcggaagatgcgcg 3913	
RESULT	4	
AA05907		
ID	AA05907 standard; DNA; 956 BP.	
XX		
AC	AA05907;	
XX		
DT	30-MAY-1996 (first entry)	
XX		
DE	Mitochondrial cytochrome c oxidase subunit III DNA.	
XX		
KW	Human; mitochondrial cytochrome C oxidase; COX; subunit I; subunit II;	
KW	subunit III; mutation; Alzheimer's disease; AD; sporadic form;	
KW	diabetes mellitus; IDDM; ss.	
XX		
CS	Homo sapiens	
XX		

FH	Key	Location/Qualifiers
FT	5'UTR	1..70
FT		/*tag= a
FT	CDS	71..856
FT		/*tag= b
FT		/product= COX C subunit III
FT	3'UTR	857..956
FT		/*tag= c
XX		
PN	W09526973-A1.	
XX		
PD	12-OCT-1995.	
XX		
PF	30-MAR-1995;	95WO-US04063.
XX		
PR	03-MAR-1995;	95US-0397808.
XX		
PR	30-MAR-1994;	94US-0219842.
XX		
PA	(GENE-) APPLIED GENETICS INC.	
XX		
PI	Davis RE, Herrnstadt C, Miller SW, Parker WD;	
XX		
DR	WPI; 1995-358577/46.	
XX		
PT	Mutant mitochondrial cytochrome C oxidase genes - useful for	
PT	generating probes for diagnosing and treating e.g. Alzheimer's	
XX	disease and new cell lines for screening for drugs.	
PS	Claim 15; Fig 3; 149pp; English.	
XX		
CC	The sequences given in AAT05905-07 encode the human mitochondrial	
CC	cytochrome C oxidase (COX) subunits I, II and III. The COX gene is	
CC	mutated in patients with Alzheimer's disease (AD) and comparison	
CC	between these sequences and those from patients known to have AD leads	
CC	to the identification of recurrent mutations. Knowledge of these	
CC	mutations allows the detection of the sporadic form of AD. Mutations	
CC	within the COX I and II genes have also been found to segregate with	
CC	diabetes mellitus. No mutations in the COX III gene was found to	
CC	significantly segregate with either AD or diabetes mellitus.	
SQ	Sequence 956 BP; 267 A; 292 C; 135 G; 262 T; 0 other;	
Query Match	31.0%; Score 27; DB 16; Length 956;	
Best Local Similarity	57.8%; Pred. No. 6.7;	
Matches	48; Conservative 0; Mismatches 35; Indels 0; Gaps 0;	
OY	4 tcgattctcgtctgctgccttgatccgatcctgcgcggtcgatgtcgcgaataacg 63	
Db	704 tcaacttcctcctacctcgtctcatccgcgaactcgatattcacttacaaccaaat 763	
OY	64 cacgtcggtttcgagatggcgcg 86	
Db	764 cacttggcttcgaagcgcgcgcg 786	
RESULT	5	
AAA67153		
ID	AAA67153 standard; DNA; 373 BP.	
AC	AAA67153;	
XX		
DT	31-OCT-2000 (first entry)	
XX		
DE	Eucalyptus grandis L.3-beta-D-glucanase DNA sequence SEQ ID NO:154.	
KW	Eucalyptus grandis; pinus radiata; Monterey pine; modification;	
KW	plant cell wall; polysaccharide; polysaccharide biosynthetic pathway;	
KW	transgenic plant; ds.	
OS	Eucalyptus grandis.	
XX		
NN	W0200022092-A2.	
NN		





```

XX      WPI: 2000-686079/67.
DR      P-PSDB: AAB26917.
XX
XX      New nucleic acid encoding bacterial nitric oxide reductase, useful e.g.
PT      for preparing transgenic organisms able to detoxify wastes
XX
XX      Example 2; Columns 31-34; 24pp; English.
XX
XX      The present sequence is the napa gene from Pseudomonas sp. strain G-179.
CC      This sequence encodes the large subunit of periplasmic nitrate reductase.
CC      This sequence can be used to create transgenic organisms that can
CC      identify toxic waste. In addition, the present sequence may be used to
CC      produce probes and primers for identification of new denitrifying
CC      bacteria.
XX
XX      Sequence 2505 BP; 557 A; 789 C; 712 G; 447 T; 0 other;
SQ
XX
XX      Query Match          29.9%; Score 26; DB 21; Length 2505;
XX      Best Local Similarity 65.5%; Pred. No. 18;
XX      Matches 38; Conservative 0; Mismatches 20; Indels 0; Gaps 0
XX
XX      Cy      30 ccgcacctctgcgcggtgcgatgtcggaacaacagcagcgtcggttcgagatgcgcga 87
XX              11 11111111 11111111 11 11 11 11 11 11 11 11 11 11
XX      Db      1575 cctgatactctgcgcgcgcgcgatgtggtgtaaaagaagagcgttaacgcaatgcgaa 1632
XX
XX      RESULT 9
XX      AAH41550/c
XX      ID      AAH41550 standard; DNA; 2765 BP.
XX
XX      AAH41550;
XX
XX      14-SEP-2001 (first entry)
XX
XX      Human Rlt1 gamma nucleotide sequence.
XX
XX      Mouse; human; combined DNA/RNA molecule; Rlt1; tumour suppressor;
KW      2-3 type zinc finger structure; cancer; diagnosis; carcinogenesis;
XX      gene therapy; ds.
XX
XX      Homo sapiens.
XX
XX      OS
XX      FS
XX      FT      Location/Qualifiers
XX      CDS      11..2116
XX              /*tag= a
XX              /product= "Rlt1 gamma"
XX
XX      WO200132859-A1.
XX
XX      PD      10-MAY-2001.
XX
XX      14-JUL-2000; 2000WO-JP04765.
XX
XX      29-OCT-1999; 99JP-0310420.
XX
XX      (MOCH ) MOCHIDA PHARM CO LTD.
XX
XX      Kominami R;
XX      PI
XX      WPI: 2001-316438/33.
XX      DR      P-PSDB: AAB99341.
XX
XX      New zinc finger protein and gene encoding it for detecting and
PT      diagnosing cancer, estimating the risk of carcinogenesis, and for gene
XX      therapy
XX
XX      Claim 40; Page 81-84; 119pp; Japanese.
XX
XX      The present invention describes a combined DNA/RNA molecule designated
CC      Rlt1, which has a 2-3 type zinc finger structure and tumour suppressor
CC      activity. Rlt1 has cytostatic activity and can be used in gene therapy.

```







XX	WO9947692-A2.
PN	
XX	
PD	23-SEP-1999.
XX	
PF	18-MAR-1999; 99WO-US05930.
XX	
PR	18-MAR-1998; 98US-0040736.
XX	
PA	(UYTO ) UNIV JOHNS HOPKINS.
PI	
XX	Rothstein JD, Lin CG, Bristol LA;
DR	WPI, 1999-580308/49.
XX	
PT	Use of glutamate transporter 2 mRNA for developing products for
XX	diagnosis or treatment of neurological disorders, e.g. amyotrophic
PT	lateral sclerosis, Huntington's, Parkinson's or Alzheimer's disease -
XX	
PS	Claim 11; Fig 4; 99pp; English.
XX	
CC	This sequence represents aberrant glutamate transporter 2
CC	(excitatory amino acid transporter 2, EAAT 2) cDNA #4, derived from
CC	human astrocyte mRNA. This sequence comprises part of the 5'
CC	untranslated region of the EAAT 2 gene and a portion of exon 8,
CC	plus exons 9-11, of the wild-type EAAT 2 cDNA (AA282111). mRNA
CC	encoding aberrant forms of EAAT 2 are abnormally spliced variants of the
CC	wild-type sequence and contain an intron sequence (AA282112) or are
CC	missing one or more exons (AA28113-228121). At least one type of
CC	aberrant EAAT 2 mRNA was detected in a majority of patients having, or
CC	suspected of having, certain neurological disorders associated with
CC	excitotoxicity which affect motor neuron function. These disorders
CC	include amyotrophic lateral sclerosis (ALS) Huntington's disease,
CC	Parkinson's disease and Alzheimer's disease. The aberrant mRNA, or cDNA
CC	derived from it, can be used as a marker for these diseases in a novel
CC	method of detection of a neurological disorder. This method comprises
CC	obtaining a biological sample from a patient, and characterizing the EAAT
CC	2 mRNA. The presence of at least one aberrant EAAT 2 mRNA in the sample
CC	is indicative of the neurological disorder in the patient. The methods
CC	can be used for detecting neurological disorders such as ALS.
CC	Huntington's disease, Parkinson's disease or Alzheimer's disease. The
CC	method can also be used for monitoring the efficiency of therapy and for
CC	identifying compounds for the treatment of neurological disorders.
CC	Recombinant vectors comprising antisense nucleic acids homologous to the
CC	aberrant sequences can be used for treating or preventing neurological
CC	disorders. The methods provide sensitive and reliable methods for
CC	detecting specific neurological disorders in living patients with minimal
CC	impact to the nervous system. In addition, the methods can serve as an
CC	indicator of predisposition for or susceptibility to the disorders.
XX	
SQ	Sequence 611 BP; 153 A; 174 C; 162 G; 122 T; 0 other;
Query Match	29.4%; Score 25.6; DB 20; Length 611;
Best Local Similarity	59.7%; Pred. No. 18;
Matches 43; Conservative 0; Mismatches 29; Indels 0; Gaps 0;	
0Y	15 gctggtctcgtcgtacatcctcgtcgcggtgtcgtatgtcgcgaacaacacgttgcgtt 74
Db	242 gctgtcacacatgcctccatcttgcagcgcgttggtgcctgcacacagagacatcagctt 301
0Y	75 cgaagatggcgcg 86
Db	302 gctgtgtgcgtgc 313

Search completed: January 10, 2002, 00:59:23  
Job time: 3176 sec





US-08-865-297-1/C  
Sequence 1, Application US/08865297  
Patent No. 6010853  
GENERAL INFORMATION:  
APPLICANT: Prasad V. S. Kanteti, Zhaohui Ao, and Stuart F. Schlozman  
TITLE OF INVENTION: The Siva Genes, No. 60108361 Genes Involved in  
TITLE OF INVENTION: CD27-Mediated Apoptosis  
NUMBER OF SEQUENCES: 12  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: LAHIVE & COCKFIELD, LLP  
STREET: 60 State Street, Suite 510  
CITY: Boston  
STATE: Massachusetts  
COUNTRY: USA  
ZIP: 02109-1875  
COMPUTER READABLE FORM:  
MEDIUM TYPE: Floppy disk  
COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: PatentIn Release #1.0, Version #1.25  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/08/865,297  
FILING DATE: 29-May-1997  
CLASSIFICATION: 435



```

PRIOR APPLICATION DATA:
  APPLICATION NUMBER:
  FILING DATE:
  ATTORNEY/AGENT INFORMATION:
    NAME: Silveri, Jean M.
  REGISTRATION NUMBER: 39,030
  REFERENCE/DOCKET NUMBER: DEN-024
  TELECOMMUNICATION INFORMATION:
    TELEPHONE: (617)227-7400
    TELEFAX: (617)227-5941
  INFORMATION FOR SEQ ID NO: 1:
    SEQUENCE CHARACTERISTICS:
      LENGTH: 885 base pairs
      TYPE: nucleic acid
      STRANDEDNESS: single
      TOPOLOGY: linear
    MOLECULE TYPE: CDNA
    FRAGMENT TYPE: Internal
  FEATURE:
    NAME/KEY: CDS
    LOCATION: 133..699
    US-08-865-297-1

```

```

Query Match          29.4%; Score 25.6; DB 3; Length 885;
Best Local Similarity 57.5%; Pred. No. 4.2;
Matches 46; Conservative 0; Mismatches 34; Indels 0; Gaps 0;

```

```

OY 7 attctctcgtgctcgtcctgcatccatccctgcgcggtgcatgctgcggaacacagc 66
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 248 AACCTCCCTGGCTGACGCGGACCTGTAGCGGGCCACGTCGGAAGGGGCGAG 189
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
OY 67 gtcggttcgagatgcccgc 86
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 188 CTCGCTTGCGCATGCGCGC 169
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||

```

```

RESULT 6
US-09-020-956-6/c
; Sequence 6, Application US/09020956
; Patent No. 6261562
; GENERAL INFORMATION:
; APPLICANT: Xu, Jiangchun
; APPLICANT: Dillon, Davin C.
; TITLE OF INVENTION: COMPOUNDS FOR IMMUNOTHERAPY OF PROSTATE CANCER AND METHODS FO
; NUMBER OF SEQUENCES: 178
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: SEED and BERRY LLP
; STREET: 6300 Columbia Center, 701 Fifth Avenue
; CITY: Seattle
; STATE: WA
; COUNTRY: USA
; ZIP: 98104
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patent Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/09/020,956
; FILING DATE: 09-FEB-1998
; CLASSIFICATION:
; ATTORNEY/AGENT INFORMATION:
; NAME: Makl, David J.
; REGISTRATION NUMBER: 31,392
; REFERENCE/DOCKET NUMBER: 210121.427C2
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (206) 622-4900
; TELEFAX: (206) 682-6031
; INFORMATION FOR SEQ ID NO: 6:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 818 base pairs
; TYPE: nucleic acid

```

```

; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: CDNA
; US-09-020-956-6

```

```

Query Match          29.2%; Score 25.4; DB 4; Length 818;
Best Local Similarity 56.6%; Pred. No. 4.8;
Matches 47; Conservative 0; Mismatches 36; Indels 0; Gaps 0;

```

```

OY 4 tgcattctcgtcgtcctgcatccatccctgcgcggtgcatgctgcggaacacag 63
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 171 TCAACTTCTCCTCACTATGCTTCATCCGCGCACTAATATTTCACTTTACATCAACAT 112
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
OY 64 cactgtggttcgagatgcccgc 86
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 111 CACTTGTGCTTCGAAGCCGCCGC 89
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||

```

```

RESULT 7
US-09-030-607-6/c
; Sequence 6, Application US/09030607
; Patent No. 6262245
; GENERAL INFORMATION:
; APPLICANT: Xu, Jiangchun
; APPLICANT: Dillon, Davin C.
; TITLE OF INVENTION: COMPOUNDS FOR IMMUNOTHERAPY OF PROSTATE CANCER AND METHODS
; NUMBER OF SEQUENCES: 224
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: SEED and BERRY LLP
; STREET: 6300 Columbia Center, 701 Fifth Avenue
; CITY: Seattle
; STATE: WA
; COUNTRY: USA
; ZIP: 98104
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patent Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/09/030,607
; FILING DATE: 25-FEB-1998
; CLASSIFICATION:
; ATTORNEY/AGENT INFORMATION:
; NAME: Makl, David J.
; REGISTRATION NUMBER: 31,392
; REFERENCE/DOCKET NUMBER: 210121.427C3
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (206) 622-4900
; TELEFAX: (206) 682-6031
; INFORMATION FOR SEQ ID NO: 6:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 818 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: CDNA
; US-09-030-607-6

```

```

Query Match          29.2%; Score 25.4; DB 4; Length 818;
Best Local Similarity 56.6%; Pred. No. 4.8;
Matches 47; Conservative 0; Mismatches 36; Indels 0; Gaps 0;

```

```

OY 4 tgcattctcgtcgtcctgcatccatccctgcgcggtgcatgctgcggaacacag 63
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 171 TCAACTTCTCCTCACTATGCTTCATCCGCGCACTAATATTTCACTTTACATCAACAT 112
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
OY 64 cactgtggttcgagatgcccgc 86
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 111 CACTTGTGCTTCGAAGCCGCCGC 89
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||

```



Query Match	29.28;	Score 25.4;	DB 4;	Length 954;
Best Local Similarity	56.68;	Pred. No. 5;		
Matches 47; Conservative	0;	Mismatches 36;	Indels 0;	Gaps 0

QY	4	tcgacatccctccctgctgctgcgcacatccctgcgcgctgcgcatatgcgcaaacag	63
Db	703	TCACATCTTCCTCACHATCTGCTTCATCCGCCACACTAATTTCACTTACATCCAAACAT	762
QY	64	caagctcgctlltcgagatgctgcgc	86
Db	763	CACCTTGGCTTGAAGCGCGCGC	785

QY	64	cacgtcgttctcagatg	ccgc	86
Db	763	CACTTTGGCTTCGAAGC	CGCCGC	785

```

RESULT 12
PCT-US95-04063-3
; Sequence 3, Application PC/TUS9504063
;
; GENERAL INFORMATION:
;
; APPLICANT: HERRNSTADT, CORINNA
;
; APPLICANT: PARKER, WILLIAM D.
;
; APPLICANT: DAVIS, ROBERT

```

Sequence 3, Application PC/TUS9504063

GENERAL INFORMATION:

APPLICANT: HERRNSTADT, CORINNA

APPLICANT: PARKER, WILLIAM D.

APPLICANT: DAVIS, ROBERT

APPLICANT: MILLER, SCOTT W.

TITLE OF INVENTION: Diagnosis, Therapy and Cellular and Animal Models for Diseases Associated with Mitochondrial

1  
2  
3  
4  
5  
6  
7  
8  
9  
10  
11  
12  
13  
14  
15  
16  
17  
18  
19  
20  
21  
22  
23  
24  
25  
26  
27  
28  
29  
30  
31  
32  
33  
34  
35  
36  
37  
38  
39  
40  
41  
42  
43  
44  
45  
46  
47  
48  
49  
50  
51  
52  
53  
54  
55  
56  
57  
58  
59  
60  
61  
62  
63  
64  
65  
66  
67  
68  
69  
70  
71  
72  
73  
74  
75  
76  
77  
78  
79  
80  
81  
82  
83  
84  
85  
86  
87  
88  
89  
90  
91  
92  
93  
94  
95  
96  
97  
98  
99  
100  
101  
102  
103  
104  
105  
106  
107  
108  
109  
110  
111  
112  
113  
114  
115  
116  
117  
118  
119  
120  
121  
122  
123  
124  
125  
126  
127  
128  
129  
130  
131  
132  
133  
134  
135  
136  
137  
138  
139  
140  
141  
142  
143  
144  
145  
146  
147  
148  
149  
150  
151  
152  
153  
154  
155  
156  
157  
158  
159  
160  
161  
162  
163  
164  
165  
166  
167  
168  
169  
170  
171  
172  
173  
174  
175  
176  
177  
178  
179  
180  
181  
182  
183  
184  
185  
186  
187  
188  
189  
190  
191  
192  
193  
194  
195  
196  
197  
198  
199  
200  
201  
202  
203  
204  
205  
206  
207  
208  
209  
210  
211  
212  
213  
214  
215  
216  
217  
218  
219  
220  
221  
222  
223  
224  
225  
226  
227  
228  
229  
230  
231  
232  
233  
234  
235  
236  
237  
238  
239  
240  
241  
242  
243  
244  
245  
246  
247  
248  
249  
250  
251  
252  
253  
254  
255  
256  
257  
258  
259  
260  
261  
262  
263  
264  
265  
266  
267  
268  
269  
270  
271  
272  
273  
274  
275  
276  
277  
278  
279  
280  
281  
282  
283  
284  
285  
286  
287  
288  
289  
290  
291  
292  
293  
294  
295  
296  
297  
298  
299  
300  
301  
302  
303  
304  
305  
306  
307  
308  
309  
310  
311  
312  
313  
314  
315  
316  
317  
318  
319  
320  
321  
322  
323  
324  
325  
326  
327  
328  
329  
330  
331  
332  
333  
334  
335  
336  
337  
338  
339  
340  
341  
342  
343  
344  
345  
346  
347  
348  
349  
350  
351  
352  
353  
354  
355  
356  
357  
358  
359  
360  
361  
362  
363  
364  
365  
366  
367  
368  
369  
370  
371  
372  
373  
374  
375  
376  
377  
378  
379  
380  
381  
382  
383  
384  
385  
386  
387  
388  
389  
390  
391  
392  
393  
394  
395  
396  
397  
398  
399  
400  
401  
402  
403  
404  
405  
406  
407  
408  
409  
410  
411  
412  
413  
414  
415  
416  
417  
418  
419  
420  
421  
422  
423  
424  
425  
426  
427  
428  
429  
430  
431  
432  
433  
434  
435  
436  
437  
438  
439  
440  
441  
442  
443  
444  
445  
446  
447  
448  
449  
450  
451  
452  
453  
454  
455  
456  
457  
458  
459  
460  
461  
462  
463  
464  
465  
466  
467  
468  
469  
470  
471  
472  
473  
474  
475  
476  
477  
478  
479  
480  
481  
482  
483  
484  
485  
486  
487  
488  
489  
490  
491  
492  
493  
494  
495  
496  
497  
498  
499  
500  
501  
502  
503  
504  
505  
506  
507  
508  
509  
510  
511  
512  
513  
514  
515  
516  
517  
518  
519  
520  
521  
522  
523  
524  
525  
526  
527  
528  
529  
530  
531  
532  
533  
534  
535  
536  
537  
538  
539  
540  
541  
542  
543  
544  
545  
546  
547  
548  
549  
550  
551  
552  
553  
554  
555  
556  
557  
558  
559  
560  
561  
562  
563  
564  
565  
566  
567  
568  
569  
570  
571  
572  
573  
574  
575  
576  
577  
578  
579  
580  
581  
582  
583  
584  
585  
586  
587  
588  
589  
590  
591  
592  
593  
594  
595  
596  
597  
598  
599  
600  
601  
602  
603  
604  
605  
606  
607  
608  
609  
610  
611  
612  
613  
614  
615  
616  
617  
618  
619  
620  
621  
622  
623  
624  
625  
626  
627  
628  
629  
630  
631  
632  
633  
634  
635  
636  
637  
638  
639  
640  
641  
642  
643  
644  
645  
646  
647  
648  
649  
650  
651  
652  
653  
654  
655  
656  
657  
658  
659  
660  
661  
662  
663  
664  
665  
666  
667  
668  
669  
670  
671  
672  
673  
674  
675  
676  
677  
678  
679  
680  
681  
682  
683  
684  
685  
686  
687  
688  
689  
690  
691  
692  
693  
694  
695  
696  
697  
698  
699  
700  
701  
702  
703  
704  
705  
706  
707  
708  
709  
710  
711  
712  
713  
714  
715  
716  
717  
718  
719  
720  
721  
722  
723  
724  
725  
726  
727  
728  
729  
730  
731  
732  
733  
734  
735  
736  
737  
738  
739  
740  
741  
742  
743  
744  
745  
746  
747  
748  
749  
750  
751  
752  
753  
754  
755  
756  
757  
758  
759  
760  
761  
762  
763  
764  
765  
766  
767  
768  
769  
770  
771  
772  
773  
774  
775  
776  
777  
778  
779  
780  
781  
782  
783  
784  
785  
786  
787  
788  
789  
790  
791  
792  
793  
794  
795  
796  
797  
798  
799  
800  
801  
802  
803  
804  
805  
806  
807  
808  
809  
810  
811  
812  
813  
814  
815  
816  
817  
818  
819  
820  
821  
822  
823  
824  
825  
826  
827  
828  
829  
830  
831  
832  
833  
834  
835  
836  
837  
838  
839  
840  
84

ADDRESS: Kenyon & Kenyon  
STREET: 1025 Connecticut Avenue, N.W.  
CITY: Washington  
STATE: DC  
COUNTRY: USA  
ZIP: 20036-5405  
COMPUTER READABLE FORM:  
MEDIUM TYPE: Floppy disk  
COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS

1  
 2  
 3  
 4  
 5  
 6  
 7  
 8  
 9  
 10  
 11  
 12  
 13  
 14  
 15  
 16  
 17  
 18  
 19  
 20  
 21  
 22  
 23  
 24  
 25  
 26  
 27  
 28  
 29  
 30  
 31  
 32  
 33  
 34  
 35  
 36  
 37  
 38  
 39  
 40  
 41  
 42  
 43  
 44  
 45  
 46  
 47  
 48  
 49  
 50  
 51  
 52  
 53  
 54  
 55  
 56  
 57  
 58  
 59  
 60  
 61  
 62  
 63  
 64  
 65  
 66  
 67  
 68  
 69  
 70  
 71  
 72  
 73  
 74  
 75  
 76  
 77  
 78  
 79  
 80  
 81  
 82  
 83  
 84  
 85  
 86  
 87  
 88  
 89  
 90  
 91  
 92  
 93  
 94  
 95  
 96  
 97  
 98  
 99  
 100  
 101  
 102  
 103  
 104  
 105  
 106  
 107  
 108  
 109  
 110  
 111  
 112  
 113  
 114  
 115  
 116  
 117  
 118  
 119  
 120  
 121  
 122  
 123  
 124  
 125  
 126  
 127  
 128  
 129  
 130  
 131  
 132  
 133  
 134  
 135  
 136  
 137  
 138  
 139  
 140  
 141  
 142  
 143  
 144  
 145  
 146  
 147  
 148  
 149  
 150  
 151  
 152  
 153  
 154  
 155  
 156  
 157  
 158  
 159  
 160  
 161  
 162  
 163  
 164  
 165  
 166  
 167  
 168  
 169  
 170  
 171  
 172  
 173  
 174  
 175  
 176  
 177  
 178  
 179  
 180  
 181  
 182  
 183  
 184  
 185  
 186  
 187  
 188  
 189  
 190  
 191  
 192  
 193  
 194  
 195  
 196  
 197  
 198  
 199  
 200  
 201  
 202  
 203  
 204  
 205  
 206  
 207  
 208  
 209  
 210  
 211  
 212  
 213  
 214  
 215  
 216  
 217  
 218  
 219  
 220  
 221  
 222  
 223  
 224  
 225  
 226  
 227  
 228  
 229  
 230  
 231  
 232  
 233  
 234  
 235  
 236  
 237  
 238  
 239  
 240  
 241  
 242  
 243  
 244  
 245  
 246  
 247  
 248  
 249  
 250  
 251  
 252  
 253  
 254  
 255  
 256  
 257  
 258  
 259  
 260  
 261  
 262  
 263  
 264  
 265  
 266  
 267  
 268  
 269  
 270  
 271  
 272  
 273  
 274  
 275  
 276  
 277  
 278  
 279  
 280  
 281  
 282  
 283  
 284  
 285  
 286  
 287  
 288  
 289  
 290  
 291  
 292  
 293  
 294  
 295  
 296  
 297  
 298  
 299  
 300  
 301  
 302  
 303  
 304  
 305  
 306  
 307  
 308  
 309  
 310  
 311  
 312  
 313  
 314  
 315  
 316  
 317  
 318  
 319  
 320  
 321  
 322  
 323  
 324  
 325  
 326  
 327  
 328  
 329  
 330  
 331  
 332  
 333  
 334  
 335  
 336  
 337  
 338  
 339  
 340  
 341  
 342  
 343  
 344  
 345  
 346  
 347  
 348  
 349  
 350  
 351  
 352  
 353  
 354  
 355  
 356  
 357  
 358  
 359  
 360  
 361  
 362  
 363  
 364  
 365  
 366  
 367  
 368  
 369  
 370  
 371  
 372  
 373  
 374  
 375  
 376  
 377  
 378  
 379  
 380  
 381  
 382  
 383  
 384  
 385  
 386  
 387  
 388  
 389  
 390  
 391  
 392  
 393  
 394  
 395  
 396  
 397  
 398  
 399  
 400  
 401  
 402  
 403  
 404  
 405  
 406  
 407  
 408  
 409  
 410  
 411  
 412  
 413  
 414  
 415  
 416  
 417  
 418  
 419  
 420  
 421  
 422  
 423  
 424  
 425  
 426  
 427  
 428  
 429  
 430  
 431  
 432  
 433  
 434  
 435  
 436  
 437  
 438  
 439  
 440  
 441  
 442  
 443  
 444  
 445  
 446  
 447  
 448  
 449  
 450  
 451  
 452  
 453  
 454  
 455  
 456  
 457  
 458  
 459  
 460  
 461  
 462  
 463  
 464  
 465  
 466  
 467  
 468  
 469  
 470  
 471  
 472  
 473  
 474  
 475  
 476  
 477  
 478  
 479  
 480  
 481  
 482  
 483  
 484  
 485  
 486  
 487  
 488  
 489  
 490  
 491  
 492  
 493  
 494  
 495  
 496  
 497  
 498  
 499  
 500  
 501  
 502  
 503  
 504  
 505  
 506  
 507  
 508  
 509  
 510  
 511  
 512  
 513  
 514  
 515  
 516  
 517  
 518  
 519  
 520  
 521  
 522  
 523  
 524  
 525

SOFTWARE: PatentIn Release #1.0, v  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: PCT/US95/0406-1995  
FILING DATE: 30-MAR-1995  
CLASSIFICATION:  
ATTORNEY/AGENT INFORMATION:  
NAME: Bonham, David B.  
REGISTRATION NUMBER: 34297  
REFERENCE/DOCKET NUMBER: 2105/7  
TELECOMMUNICATION INFORMATION:









```

; Sequence 2, Application US/08865297
; Patent No. 6010853
; GENERAL INFORMATION:
; APPLICANT: Prasad V.S. Kanleti, Zhaochui Ao, and Stuart F.
; TITLE OF INVENTION: The Siva Genes, No. 6010853el Genes Involved in
; TITLE OF INVENTION: CD27-Mediated Apoptosis
; NUMBER OF SEQUENCES: 12
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: LAHIVE & COCKFIELD, LLP
; STREET: 60 State Street, Suite 510
; CITY: Boston
; STATE: Massachusetts
; COUNTRY: USA
; ZIP: 02109-1875
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/865,297
; FILING DATE: 29-MAY-1997
; CLASSIFICATION: 435
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER:
; FILING DATE:
; ATTORNEY/AGENT INFORMATION:
; NAME: Silveri, Jean M.
; REGISTRATION NUMBER: 39,030
; REFERENCE/DOCKET NUMBER: DFN-024
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (617)227-7400
; TELEFAX: (617)227-5941
; INFORMATION FOR SEQ ID NO: 2:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 189 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: protein
; US-08-865-297-2

```

```

alignment_scores:
  Quality: 66.00      Length: 27
  Ratio: 3.000      Gaps: 0
  Percent Similarity: 81.481      Percent Identity: 40.741

```

## alignment\_block:

```
US-09-712-768-7/rev x US-08-865-297-2 ..
```

```
Align seg 1/1 to: US-08-865-297-2 from: 1 to: 189
```

```

86 GCGGCATCTCGAAGACGACGTGTTTGGCAGCATCGCACCAGCGAG 37
|||||:|||||:|||||:|||||:|||||:|||||:|||||:
13 AlaAlaMetProLysArgSerCysProPheAlaAspValAlaProLeuGl 29
:::|||||:|||||:|||||:|||||:|||||:|||||:
36 GATCGCGATCGACGACGACGACGACGACGACGACGACGACGACGAC 6
:::|||||:|||||:|||||:|||||:|||||:|||||:
29 nleulysValArgValSerGlnArgGluLeu 39
:::|||||:|||||:|||||:|||||:|||||:|||||:

```

```
seq_name: /cgn2_6/ptodata/2/1aa/6B.COMB.pep:US-09-097-889-19
```

## seq\_documentation\_block:

```

; Sequence 19, Application US/09097889
; Patent No. 6218117
; GENERAL INFORMATION:
; APPLICANT: Herrstadt, Corrina
; APPLICANT: Ghosh, Soumitra S.
; APPLICANT: Davis, Robert E.
; TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR IDENTIFYING
; TITLE OF INVENTION: AGENTS THAT QUANTITATIVELY ALTER DETECTABLE
; TITLE OF INVENTION: EXTRAMITOCHONDRIAL DNA: MITOCHONDRIAL DNA RATIOS

```

```

; NUMBER OF SEQUENCES: 26
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: SEED and BERRY LLP
; STREET: 6300 Columbia Center, 701 Fifth Avenue
; CITY: Seattle
; STATE: Washington
; COUNTRY: USA
; ZIP: 98104
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/09/097,889
; FILING DATE: 15-JUN-1998
; CLASSIFICATION: 435
; ATTORNEY/AGENT INFORMATION:
; NAME: Rosenman Ph.D., Stephen J.
; REGISTRATION NUMBER: 43,058
; REFERENCE/DOCKET NUMBER: 660088.417
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (206) 622-4900
; TELEFAX: (206) 682-6031
; INFORMATION FOR SEQ ID NO: 19:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 261 amino acids
; TYPE: amino acid
; STRANDEDNESS:
; TOPOLOGY: linear
; US-09-097-889-19

```

```

alignment_scores:
  Quality: 66.00      Length: 28
  Ratio: 3.300      Gaps: 0
  Percent Similarity: 71.429      Percent Identity: 50.000

```

## alignment\_block:

```
US-09-712-768-7 x US-09-097-889-19 ..
```

```
Align seg 1/1 to: US-09-097-889-19 from: 1 to: 261
```

```

4 TCGATCTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCGGCGTGCAGATGC 53
||| ||||| :|||:|||||:|||||:|||||:|||||:|||||:
212 SerThrPheLeuThrIleCysPheIleArgGlnLeuMetPheHisPheTh 228
:::|||||:|||||:|||||:|||||:|||||:|||||:
54 GGCANAACAGCAGCGTGGTTTCGAGATGCCGCGCA 87
:::|||||:|||||:|||||:|||||:|||||:|||||:
228 rSerLysHisHisPheGlyPheGlnAlaAlaAla 239
:::|||||:|||||:|||||:|||||:|||||:|||||:

```

```
seq_name: /cgn2_6/ptodata/2/1aa/6B.COMB.pep:US-08-924-345-2
```

## seq\_documentation\_block:

```

; Sequence 2, Application US/08924345
; Patent No. 6224878
; GENERAL INFORMATION:
; APPLICANT: LEUNG-TACK Patricia
; APPLICANT: LEGASTRELOIS Isabelle, Christine, Marie-Andree
; APPLICANT: AUDONNET Jean-Christophe, Francis
; APPLICANT: RIVIERE Michel, Emile, Albert
; TITLE OF INVENTION: Mutants and vaccines of the Infectious
; TITLE OF INVENTION: Bovine Rhinotracheitis virus
; NUMBER OF SEQUENCES: 7
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: LARSON AND TAYLOR
; STREET: 727 SOUTH TWENTY-THIRD STREET
; CITY: ARLINGTON
; STATE: VIRGINIA
; COUNTRY: USA
; ZIP: 22202
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk

```



```

;
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.25 (EPO)
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/924,345
; FILING DATE: 04-SEP-1997
; CLASSIFICATION: 424
; PRIORITY APPLICATION DATA:
; APPLICATION NUMBER: 08/199,172
; FILING DATE: 09-AUG-1994
; APPLICATION NUMBER: FR 92 07930
; FILING DATE: 26-JUN-1992
; ATTORNEY/AGENT INFORMATION:
; NAME: SARRO, THOMAS P.
; REGISTRATION NUMBER: 19396
; REFERENCE/DOCKET NUMBER: XI
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (703) 920-7200
; TELEFAX: (703) 892-8428
; INFORMATION FOR SEQ ID NO: 2:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 380 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: protein
; US-08-924-345-2

```

```

alignment_scores:
  Quality: 55.50      Length: 28
  Ratio: 3.265        Gaps: 1
  Percent Similarity: 60.714  Percent Identity: 50.000

```

```

alignment_block:
  US-09-712-768-7/rev x US-08-924-345-2  ..
  Align seg 1/1 to: US-08-924-345-2 from: 1 to: 380

```

```

70 CGAGCTGCTGTTTCCGACATCGCAGCGC..... 41
||||| :||||| |||:|||||
321 ArgAgaArgLeuLeuProCysSerArgValaIlyIlySProArgThrcy 337
40 .GCAGATCGGATCAGGACAGCAGCAGAGAGA 8
||||:||||| |||:|||||
337 sAlaAlaCysclYserGlyThrcySAlaGlyArg 348

```

```

seq_name: /cgn2.6/ptodata/2/1aa/5B_COMB.pep:US-09-025-819-2

```

```

seq_documentation_block:
; Sequence 2, Application US/09025819

```

```

; Patent No. 6225097
; GENERAL INFORMATION:
; APPLICANT: Odata, Shusei
; APPLICANT: Nishino, Tokuzo
; APPLICANT: Koyama, Tanetoshi
; APPLICANT: Sato, Yoshihiro
; TITLE OF INVENTION: DECAPRENYL DIPHOSPHATE SYNTHETASE GENE
; NUMBER OF SEQUENCES: 31
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: KENYON & KENYON
; STREET: 1500 K Street, N.W.
; CITY: Washington
; STATE: DC
; COUNTRY: USA
; ZIP: 20005
; COMPUTER READABLE FORM:
; MEDIUM TYPE: floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/09/025,819
; FILING DATE: 19-FEB-1998

```

```

;
; CLASSIFICATION: 435
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: JP 251675
; FILING DATE: 17-SEP-1997
; ATTORNEY/AGENT INFORMATION:
; NAME: Khalilian, Hourl
; REGISTRATION NUMBER: 39,546
; REFERENCE/DOCKET NUMBER: 10235/2
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 202-220-4200
; TELEFAX: 202-220-4201
; INFORMATION FOR SEQ ID NO: 2:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 331 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: protein
; US-09-025-819-2

```

```

alignment_scores:
  Quality: 54.00      Length: 29
  Ratio: 2.700        Gaps: 2
  Percent Similarity: 68.966  Percent Identity: 51.724

```

```

alignment_block:
  US-09-712-768-7/rev x US-09-025-819-2  ..
  Align seg 1/1 to: US-09-025-819-2 from: 1 to: 331

```

```

85 CGGCCATCTGAAACGACGAGTGT...GTTTCCGACATCGCAGCGGC 39
||||| |||:||||| ||| ||| |||
252 ArgProArgArgArgGlyArgAlaArgLeuLeuGlyThhHisArgPr 268
38 AGGA...TGGCGATCAGCAGACGACGAGCAGAGATCG 5
||| :||| ||| ||| ||| |||
268 cGlyProAlaGlyArgGlyArgProGlyHisArgAla 280

```

```

seq_name: /cgn2.6/ptodata/2/1aa/backfill.es1.pep:5183745-2

```

```

seq_documentation_block:
; Patent No. 5183745

```

```

; APPLICANT: DANCHIN, ANTOINE;GLASER, PHILIPPE;KRIN, EVELYN;
; BARZU, OCTAVIEN;LADANT, DANIEL;ULLMAN, AGNES
; TITLE OF INVENTION: ADENYL CYCLASE DERIVATIVES AND THEIR
; BIOLOGICAL USES
; NUMBER OF SEQUENCES: 13
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/426,541
; FILING DATE: 25-OCT-1989
; SEQ ID NO: 2:
; LENGTH: 1489
; 5183745-2

```

```

alignment_scores:
  Quality: 53.00      Length: 18
  Ratio: 4.077        Gaps: 0
  Percent Similarity: 72.222  Percent Identity: 50.000

```

```

alignment_block:
  US-09-712-768-7/rev x 5183745-2  ..
  Align seg 1/1 to: 5183745-2 from: 1 to: 1489

```

```

83 GCCATCTGAAACGACGAGTGTGTTTCCGACATCGCAGCGCGCAGAT 34
||||| :||||| ||| :|||||
1413 AlaMetArgArgProThrcyScySAlaAlaArgValaIaProThrcySle 1429
33 CGCG 30
: |||
1429 uArg 1430

```

```
seq_name: /cgn2_6/ptodata/2/1aa/6A_COMB.pep:US-08-996-139-2
seq_documentation_block:
; Sequence 2, Application US/08996139
; Patent No. 6017729
; GENERAL INFORMATION:
; APPLICANT: Anderson, Dirk M.
; APPLICANT: Galibert, Laurent
; APPLICANT: Maraskovsky, Eugene
; TITLE OF INVENTION: Receptor Activator of NF-kappaB
; NUMBER OF SEQUENCES: 19
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Immunex Corporation, Law Department
; STREET: 51 University Street
; CITY: Seattle
; STATE: WA
; COUNTRY: USA
; ZIP: 98101
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: Apple Power Macintosh
; OPERATING SYSTEM: Apple Operating System 7.5.5
; SOFTWARE: Microsoft Word for Power Macintosh 6.0.1
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/996,139
; FILING DATE: 22 DECEMBER 1997
; CLASSIFICATION:
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: USSN 60/064,671
; FILING DATE: 14 OCTOBER 1997
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: USSN 08/813,509
; FILING DATE: 07 MARCH 1997
; CLASSIFICATION:
; APPLICATION NUMBER: USSN 08/772,330
; FILING DATE: 23 DECEMBER 1996
; ATTORNEY/AGENT INFORMATION:
; NAME: Perkins, Patricia Anne
; REGISTRATION NUMBER: 34,693
; REFERENCE/DOCKET NUMBER: 2851-A
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (206)587-0430
; TELEFAX: (206)233-0644
; INFORMATION FOR SEQ ID NO: 2:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 591 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: protein
; US-08-996-139-2

alignment_scores:
Quality: 51.50 Length: 19
Ratio: 3.219 Gaps: 1
Percent Similarity: 84.211 Percent Identity: 63.158

alignment_block:
US-09-712-768-7/rev x US-08-996-139-2 ..
Align seg 1/1 to: US-08-996-139-2 from: 1 to: 591

58 TTGCGCATGCGACGCGAGATGCGATCAGGC...AGACGACGAG 12
|||||:||||| |||||||:||||| |||:|
476 LeupProSerSerAlaArgAlaGlyAlaGlySerGlySerSerProGlyI 492
11 AAGATCG 5
|:|:|:|
492 yGlnSer 494

seq_name: /cgn2_6/ptodata/2/1aa/6B_COMB.pep:US-08-995-659-2
seq_documentation_block:
```

```
; Sequence 2, Application US/08995659
; Patent No. 6242213
; GENERAL INFORMATION:
; APPLICANT: Anderson, Dirk M.
; APPLICANT: Galibert, Laurent
; APPLICANT: Maraskovsky, Eugene
; TITLE OF INVENTION: Ligand for Receptor Activator of NF-kappaB
; NUMBER OF SEQUENCES: 19
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Immunex Corporation, Law Department
; STREET: 51 University Street
; CITY: Seattle
; STATE: WA
; COUNTRY: USA
; ZIP: 98101
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: Apple Power Macintosh
; OPERATING SYSTEM: Apple Operating System 7.5.5
; SOFTWARE: Microsoft Word for Power Macintosh 6.0.1
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/995,659
; FILING DATE: 22 DECEMBER 1997
; CLASSIFICATION:
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: USSN 60/064,671
; FILING DATE: 14 OCTOBER 1997
; CLASSIFICATION:
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: USSN 08/813,509
; FILING DATE: 07 MARCH 1997
; CLASSIFICATION:
; APPLICATION NUMBER: USSN 08/772,330
; FILING DATE: 23 DECEMBER 1996
; ATTORNEY/AGENT INFORMATION:
; NAME: Perkins, Patricia Anne
; REGISTRATION NUMBER: 34,693
; REFERENCE/DOCKET NUMBER: 2852-A
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (206)587-0430
; TELEFAX: (206)233-0644
; INFORMATION FOR SEQ ID NO: 2:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 591 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: protein
; US-08-995-659-2

alignment_scores:
Quality: 51.50 Length: 19
Ratio: 3.219 Gaps: 1
Percent Similarity: 84.211 Percent Identity: 63.158

alignment_block:
US-09-712-768-7/rev x US-08-995-659-2 ..
Align seg 1/1 to: US-08-995-659-2 from: 1 to: 591

58 TTGCGCATGCGACGCGAGATGCGATCAGGC...AGACGACGAG 12
|||||:||||| |||||||:||||| |||:|
476 LeupProSerSerAlaArgAlaGlyAlaGlySerGlySerSerProGlyI 492
11 AAGATCG 5
|:|:|:|
492 yGlnSer 494

seq_name: /cgn2_6/ptodata/2/1aa/6B_COMB.pep:US-09-215-649A-2
seq_documentation_block:
```

```

Sequence 2, Application US/09215649A
Patent No. 6271349
GENERAL INFORMATION:
APPLICANT: Anderson, Dirk M.
Galibert, Laurent
Maraskovsky, Eugene
TITLE OF INVENTION: Receptor Activator of NF-kappaB
NUMBER OF SEQUENCES: 19
CORRESPONDENCE ADDRESS:
ADDRESSEE: Immunex Corporation, Law Department
STREET: 51 University Street
CITY: Seattle
STATE: WA
COUNTRY: USA
ZIP: 98101
COMPUTER READABLE FORM:
MEDIUM TYPE: floppy disk
COMPUTER: Apple Power Macintosh
OPERATING SYSTEM: Apple Operating System 7.5.5
SOFTWARE: Microsoft Word for Power Macintosh 6.0.1
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/09/215,649A
FILING DATE: 17-Dec-1998
CLASSIFICATION: <unknown>
PRIOR APPLICATION DATA:
APPLICATION NUMBER: 08/996,139
FILING DATE: <unknown>
APPLICATION NUMBER: USSN 08/813,509
FILING DATE: 07 MARCH 1997
APPLICATION NUMBER: USSN 08/772,330
FILING DATE: 23 DECEMBER 1996
ATTORNEY/AGENT INFORMATION:
NAME: Perkins, Patricia Anne
REGISTRATION NUMBER: 34,693
REFERENCE/DOCKET NUMBER: 2851-A
TELECOMMUNICATION INFORMATION:
TELEPHONE: (206)587-0430
TELEFAX: (206)233-0644
INFORMATION FOR SEQ ID NO: 2:
SEQUENCE CHARACTERISTICS:
LENGTH: 591 amino acids
TYPE: amino acid
TOPOLOGY: linear
MOLECULE TYPE: protein
SEQUENCE DESCRIPTION: SEQ ID NO: 2:
US-09-215-649A-2

alignment_scores:
Quality: 51.50 Length: 19
Ratio: 3.219 Gaps: 1
Percent Similarity: 84.211 Percent Identity: 63.158

alignment_block:
US-09-712-768-7/rev x US-09-215-649A-2 ..
Align seg 1/1 to: US-09-215-649A-2 from: 1 to: 591
58 TTGCCGACATCGCACCGCGAGATGCGGATCAGGC...AGACGAGCAGG 12
|||||:||||| |||||||:||||| |||:|
476 LeuProSerSerAlaArgAlaGlyAlaGlySerGlySerSerProGlyGI 492
11 AAGATCG 5
|:|:|
492 yGlnSer 494

seq_name: /cgn2_6/ptodata/2/1aa/6A_COMB.pep:US-08-996-139-6
seq_documentation_block:
; Sequence 6, Application US/08996139
; Patent No. 6017729
; GENERAL INFORMATION:
; APPLICANT: Anderson, Dirk M.

```

```

APPLICANT: Galibert, Laurent
APPLICANT: Maraskovsky, Eugene
TITLE OF INVENTION: Receptor Activator of NF-kappaB
NUMBER OF SEQUENCES: 19
CORRESPONDENCE ADDRESS:
ADDRESSEE: Immunex Corporation, Law Department
STREET: 51 University Street
CITY: Seattle
STATE: WA
COUNTRY: USA
ZIP: 98101
COMPUTER READABLE FORM:
MEDIUM TYPE: floppy disk
COMPUTER: Apple Power Macintosh
OPERATING SYSTEM: Apple Operating System 7.5.5
SOFTWARE: Microsoft Word for Power Macintosh 6.0.1
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/996,139
FILING DATE: 22 DECEMBER 1997
CLASSIFICATION:
PRIOR APPLICATION DATA:
APPLICATION NUMBER: USSN 60/064,671
FILING DATE: 14 OCTOBER 1997
PRIOR APPLICATION DATA:
APPLICATION NUMBER: USSN 08/813,509
FILING DATE: 07 MARCH 1997
PRIOR APPLICATION DATA:
APPLICATION NUMBER: USSN 08/772,330
FILING DATE: 23 DECEMBER 1996
ATTORNEY/AGENT INFORMATION:
NAME: Perkins, Patricia Anne
REGISTRATION NUMBER: 34,693
REFERENCE/DOCKET NUMBER: 2851-A
TELECOMMUNICATION INFORMATION:
TELEPHONE: (206)587-0430
TELEFAX: (206)233-0644
INFORMATION FOR SEQ ID NO: 6:
SEQUENCE CHARACTERISTICS:
LENGTH: 616 amino acids
TYPE: amino acid
TOPOLOGY: linear
MOLECULE TYPE: protein
US-08-996-139-6

alignment_scores:
Quality: 51.50 Length: 19
Ratio: 3.219 Gaps: 1
Percent Similarity: 84.211 Percent Identity: 63.158

alignment_block:
US-09-712-768-7/rev x US-08-996-139-6 ..
Align seg 1/1 to: US-08-996-139-6 from: 1 to: 616
58 TTGCCGACATCGCACCGCGAGATGCGGATCAGGC...AGACGAGCAGG 12
|||||:||||| |||||||:||||| |||:|
501 LeuProSerSerAlaArgAlaGlyAlaGlySerGlySerSerProGlyGI 517
11 AAGATCG 5
|:|:|
517 yGlnSer 519

seq_name: /cgn2_6/ptodata/2/1aa/6B_COMB.pep:US-08-995-659-6
seq_documentation_block:
; Sequence 6, Application US/08995659
; Patent No. 6242213
; GENERAL INFORMATION:
; APPLICANT: Anderson, Dirk M.
; APPLICANT: Galibert, Laurent
; APPLICANT: Maraskovsky, Eugene
; TITLE OF INVENTION: Ligand for Receptor Activator of NF-kappaB

```



NAME: Murphy, Jr., Gerald M

REGISTRATION NUMBER: 28,977  
REFERENCE/DOCKET NUMBER: 1781-105P  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: (703) 205-8000  
TELEFAX: (703) 205-8050  
TELEX: 248345  
INFORMATION FOR SEQ ID NO: 4:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 1019 amino acids  
TYPE: amino acid  
TOPOLOGY: linear  
MOLECULE TYPE: protein  
US-08-596-405-4

alignment\_scores:  
Quality: 50.00 Length: 35  
Ratio: 2.941 Gaps: 1  
Percent Similarity: 48.571 Percent Identity: 31.429

alignment\_block:  
US-09-712-768-7 x US-08-596-405-4 ..

Align seg 1/1 to: US-08-596-405-4 from: 1 to: 1019

17 TGGTCTGCTGATCCGATCCTGCGGCTGCGATGCGCA..... 58  
|||||:|:| |||||  
741 TrpSerGlyArgProAlaSerCysIleProValCysGlyArgSerAsp 757  
59 .....AACAGCACGTCGGTTTCAGAT 80  
757 rProArgSerProPheIleTrpAsnGlyAsnSerThrGluIleGlyI 774  
81 GGCGG 85  
|||||  
774 rPPro 775

GenCore version 4.5  
Copyright (c) 1993 - 2000 Compugen Ltd.

OM protein - protein search, using sw model

Run on: January 10, 2002, 01:56:27 ; Search time 31.32 seconds  
(without alignments)  
89.872 Million cell updates/sec

Title: US-09-712-768-6  
Perfect score: 203  
Sequence: 1 IVHGDRKRTAIGLAIGLWFTLCQAVEYEIVHTE 38

Scoring table: BLOSUM62  
Gapop 10.0 , Gapext 0.5

Searched: 522463 seqs, 74073290 residues  
Total number of hits satisfying chosen parameters: 522463

Minimum DB seq length: 0  
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%  
Maximum Match 100%  
Listing first 45 summaries

Database : A.GeneSeq\_1101.\*  
1: /SIDS2/gcgdata/geneseq/geneseq/AA1980.DAT:\*  
2: /SIDS2/gcgdata/geneseq/geneseq/AA1981.DAT:\*  
3: /SIDS2/gcgdata/geneseq/geneseq/AA1982.DAT:\*  
4: /SIDS2/gcgdata/geneseq/geneseq/AA1983.DAT:\*  
5: /SIDS2/gcgdata/geneseq/geneseq/AA1984.DAT:\*  
6: /SIDS2/gcgdata/geneseq/geneseq/AA1985.DAT:\*  
7: /SIDS2/gcgdata/geneseq/geneseq/AA1986.DAT:\*  
8: /SIDS2/gcgdata/geneseq/geneseq/AA1987.DAT:\*  
9: /SIDS2/gcgdata/geneseq/geneseq/AA1988.DAT:\*  
10: /SIDS2/gcgdata/geneseq/geneseq/AA1989.DAT:\*  
11: /SIDS2/gcgdata/geneseq/geneseq/AA1990.DAT:\*  
12: /SIDS2/gcgdata/geneseq/geneseq/AA1991.DAT:\*  
13: /SIDS2/gcgdata/geneseq/geneseq/AA1992.DAT:\*  
14: /SIDS2/gcgdata/geneseq/geneseq/AA1993.DAT:\*  
15: /SIDS2/gcgdata/geneseq/geneseq/AA1994.DAT:\*  
16: /SIDS2/gcgdata/geneseq/geneseq/AA1995.DAT:\*  
17: /SIDS2/gcgdata/geneseq/geneseq/AA1996.DAT:\*  
18: /SIDS2/gcgdata/geneseq/geneseq/AA1997.DAT:\*  
19: /SIDS2/gcgdata/geneseq/geneseq/AA1998.DAT:\*  
20: /SIDS2/gcgdata/geneseq/geneseq/AA1999.DAT:\*  
21: /SIDS2/gcgdata/geneseq/geneseq/AA2000.DAT:\*  
22: /SIDS2/gcgdata/geneseq/geneseq/AA2001.DAT:\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

## SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	203	100.0	38	22	AA97752
2	62	30.5	93	22	AAB63840
3	62	30.5	133	21	AAB29910
4	62	30.5	203	22	AAB75604
5	62	30.5	455	20	AA935610
6	60	29.6	133	21	AAB29909
7	60	29.6	172	21	AAB36196
8	60	29.6	203	22	AAB75603
9	58	28.6	612	22	AAB93232
10	58	28.6	747	20	AA941758
11	58	28.6	747	21	AAB44314

12	58	28.6	747	22	AA012384	Human PRO1012 poly
13	58	28.6	768	22	AAB92743	Human protein sequ
14	58	28.6	793	22	AAB93204	Human protein sequ
15	58	28.6	793	22	AAB95309	Human protein sequ
16	58	28.6	793	22	AAB71913	Human ISOM-5. Hom
17	55	27.1	78	22	AAB75551	Human secreted pro
18	55	27.1	93	22	AA675934	Human colon cancer
19	55	27.1	847	22	AA692548	C glutamicum prote
20	54.5	26.8	131	22	AA674197	OmPA-J domain fusi
21	54.5	26.8	131	22	AA670767	Expression plasmid
22	54.5	26.8	131	22	AA672018	E. coli OmPA-J dom
23	54.5	26.8	399	21	AAB11395	E. coli expression
24	54.5	26.8	399	22	AAB74196	OmPA-DnaJ fusion p
25	54.5	26.8	399	22	AA670766	Expression plasmid
26	54.5	26.8	399	22	AA672017	E. coli OmPA-DnaJ
27	54	26.6	48	22	AA690908	C glutamicum prote
28	53.5	26.4	131	21	AAB11396	E. coli expression
29	53.5	26.4	326	21	AA646264	Arabidopsis thalia
30	53.5	26.4	336	21	AA646263	Arabidopsis thalia
31	53.5	26.4	1147	21	AA630843	Arabidopsis thalia
32	53.5	26.4	1215	21	AA630842	Arabidopsis thalia
33	53.5	26.4	1355	21	AA630841	Arabidopsis thalia
34	53	26.1	117	20	AA670366	Fragment of human
35	53	26.1	216	22	AA670366	Human cancer relat
36	53	26.1	259	20	AA670366	Human SBHNC50 pro
37	53	26.1	276	19	AA670321	Secreted protein C
38	53	26.1	488	21	AA696758	A. terreus ORF16 M
39	53	26.1	1816	21	AA695440	Caenorhabditis ele
40	51.5	25.4	192	20	AA631662	Staphylococcus aur
41	51	25.1	376	22	AA631662	G protein-coupled
42	50.5	24.9	265	20	AA609510	Human M6b1 protein
43	50.5	24.9	870	20	AA635233	C. pneumoniae prote
44	50	24.6	205	22	AA692148	C glutamicum prote
45	50	24.6	217	22	AA679574	Corynebacterium g1

## ALIGNMENTS

RESULT 1  
ID AA97752 standard; Protein: 38 AA.

AC AA97752;

DT 06-AUG-2001 (first entry)

DE G. oxydans cytochrome C oxidase (COIII) protein sequence fragment.

XX Cytochrome C oxidase; COI; COII; COIII; COI complex; D-sorbitol;

KW oxidative fermentation; electron transfer; respiratory chain; L-sorbose;

KW 2-keto-L-gluconic acid; 2KGA production; aldehyde production;

KW carboxylic acid production; ketone production.

XX Gluconobacter oxydans.

XX EPI103603-A2.

XX 30-MAY-2001.

XX 14-NOV-2000; 2000EP-0124785.

XX 17-NOV-1999; 99EP-0122842.

XX (HOFF) HOFFMANN LA ROCHE & CO AG F.

XX Asakura A, Hoshino T, Shinjoh M;

XX WPI: 2001-357953/38.

XX N-PSDB; AAA91491.

XX New cytochrome c oxidase complex having cytochrome c oxidase activity

XX from Gluconobacter oxydans DSM 4025, useful in mediating electron

PT transfer in respiratory chain or producing 2-keto-L-gulononic acid from  
 PT L-sorbose or D-sorbitol  
 XX  
 PS Claim 6; Page 25; 42pp; English.

CC This sequence represents a fragment of the Gluconobacter oxydans  
 CC cytochrome c oxidase, COIIT. The invention relates to a cytochrome c  
 CC oxidase COI complex. The COI complex is useful in improving oxidative  
 CC fermentation and is an essential component mediating electron transfer  
 CC in the respiratory chain. The recombinant microorganism and the  
 CC cytochrome c oxidase may be used in the genetic preparation of a  
 CC recombinant COI complex and in the production of 2-keto-L-gulononic acid  
 CC (2KGA) from L-sorbose or D-sorbitol in a culture medium. The COI is also  
 CC useful as a terminal oxidase, oxidizing cytochrome c, an electron  
 CC acceptor from an enzyme belonging to dehydrogenase for the production of  
 CC aldehydes, carboxylic acids and ketones from alcohols and aldehydes,  
 CC especially the production of 2KGA from L-sorbose or D-sorbitol.

XX  
 SQ Sequence 38 AA;

Query Match 100.0%; Score 203; DB 22; Length 38;  
 Best Local Similarity 100.0%; Pred. No. 2.1e-23;  
 Matches 38; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 IVHGDRKRTAIGLAIGLWIFLCAQYEEYEVITE 38  
 |||||  
 Db 1 lvhgdrkrtalglalglwiflclgqeyeyelwhte 38

RESULT 2  
 AAB63840  
 ID AAB63840 standard; Protein; 93 AA.

XX AAB63840;

XX 26-MAR-2001 (first entry)

XX Human prostate cancer associated antigen protein sequence SEQ ID NO:1202.

XX Human: breast cancer; gastric cancer; prostate cancer; diagnosis;

XX cancer associated antigen; cytostatic; cancer vaccine.

XX Homo sapiens.

XX WO200073801-A2.

XX 07-DEC-2000.

XX 26-MAY-2000; 2000WO-US14749.

XX 28-MAY-1999; 99US-0136526.

XX 10-SEP-1999; 99US-0153454.

XX (LUDW-) LUDWIG INST CANCER RES.

XX Obata Y;

XX WPI; 2001-025274/03.

XX Nucleic acids encoding breast, gastric and prostate cancer associated  
 PT antigen precursors, useful for diagnosing and treating a condition  
 PT characterized by expression of an abnormal amount of a protein, e.g.  
 PT cancer

XX Example 1; Page 741; 799pp; English.

XX AAF22422 to AAF22626, AAF22627 to AAF22773 and AAF22774 to AAF23014  
 CC represent nucleotide sequences encoding human breast, gastric and  
 CC prostate cancer associated antigen precursors (CAAP) respectively.  
 CC AAB63332 to AAB63467, AAB63468 to AAB63721 and AAB63722 to AAB63970  
 CC represent human breast, gastric and prostate CAAP protein sequence  
 CC respectively. CAAPs have cytostatic activity and can be used in the

CC production of cancer vaccines. The human CAAP proteins, peptides, nucleic  
 CC acids or anti-CAAP antibodies are useful for diagnosing and treating a  
 CC condition characterised by expression of an abnormal amount of a protein,  
 CC e.g. cancer.

XX  
 SQ Sequence 93 AA;

Query Match 30.5%; Score 62; DB 22; Length 93;  
 Best Local Similarity 42.4%; Pred. No. 0.088;  
 Matches 14; Conservative 5; Mismatches 14; Indels 0; Gaps 0;

OY 1 IVHGDRKRTAIGLAIGLWIFLCAQYEEYEVITE 33  
 ::|||  
 Db 5 llnnrrqigallitlllglyftllqaseyfe 37

RESULT 3  
 AAB29910  
 ID AAB29910 standard; Protein; 133 AA.

XX AAB29910;

XX 09-FEB-2001 (first entry)

XX Human secreted protein BLAST search protein SEQ ID NO: 168.

XX Cytostatic; immunosuppressive; nootropic; neuroprotective; antiviral;  
 KW anti-allergic; hepatotropic; antidiabetic; anti-inflammatory; antitumor;  
 KW vulnerrary; anticonvulsant; antibacterial; antifungal; antiparasitic;  
 KW cardiant; gene therapy; cancer; immune disorder; cardiovascular disorder;  
 KW neurological disease; infection; human; secreted protein.

XX Homo sapiens.

XX WO200061779-A1.

XX 19-OCT-2000.

XX 06-APR-2000; 2000WO-US09068.

XX 09-APR-1999; 99US-0128699.

XX 20-JAN-2000; 2000US-0177050.

XX (HUMA-) HUMAN GENOME SCI INC.

XX Rosen CA, Ruben SM, Komatsoulis G;

XX WPI; 2000-647424/52.

XX Isolated nucleic acid molecule encoding a human secreted protein 1s  
 PT used in preventing, treating or ameliorating a medical condition -  
 PT disclosure; Page 488-489; 495pp; English.

XX The invention relates to the isolation of genes AAG63410-C63458 encoding  
 CC the human secreted proteins AAB29802-B29850. This sequence represents a  
 CC fragment of the protein encoded by the gene given in the descriptor  
 CC line. The sequence is used as a query sequence for doing BLASTX searches  
 CC to determine homologous sequence to the protein. The genes and proteins  
 CC are useful for preventing, ameliorating or treating medical conditions,  
 CC e.g. by protein or gene therapy. The genes are isolated from a range of  
 CC human tissues disclosed in the specification. The nucleic acids,  
 CC proteins, antibodies and (ant)agonists are useful in the diagnosis,  
 CC treatment and prevention of: (a) cancer, e.g. breast and ovarian cancer,  
 CC and other cancers of the adrenal gland, bone, bone marrow, breast,  
 CC gastrointestinal tract, liver, lung, or urogenital; (b) immune disorders  
 CC e.g. Addison's disease, allergies, autoimmune haemolytic anaemia,  
 CC autoimmune thyroiditis, diabetes mellitus, Crohn's disease, multiple  
 CC sclerosis, rheumatoid arthritis and ulcerative colitis; (c)  
 CC cardiovascular disorders such as myocardial ischaemias; (d) wound  
 CC healing; (e) neurological diseases e.g. cerebral anoxia and epilepsy; and  
 CC (f) infectious diseases such as viral, bacterial, fungal and parasitic



CC infections.  
XX  
SQ Sequence 133 AA:

Query Match 30.5%; Score 62; DB 21; Length 133;  
Best Local Similarity 42.4%; Pred. No. 0.13;  
Matches 14; Conservative 5; Mismatches 14; Indels 0; Gaps 0;

OY 1 IVHGDRKRTAIGLAIAIGLGIFTLCQAVEYVE 33  
Db 23 Ilenmrnglqallllllllylllgaseyfe 55

RESULT 4  
AAB75604  
ID AAB75604 standard; Protein: 203 AA.

AC AAB75604;

DT 06-APR-2001 (first entry)

DE Human secreted protein sequence encoded by gene 46 spq ID NO:158.

XX Human: secreted protein; immunosuppressive; antiarthritic; antirheumatic;  
KM antiproliferative; cytostatic; cardiant; vasotropic; cerebroprotective;  
KM neurotropic; neuroprotective; antibacterial; virucide; fungicide;  
KM ophthalmological; vulnery; autoimmune disease; cardiovascular disorder;  
KM hyperproliferative disorders; cerebrovascular disorder; wound healing;  
KM nervous system disorder; ocular disorder; skin aging; chemotaxis;  
KM food additive.

XX Homo sapiens.

OS NC200077026-A1.

PN 21-DEC-2000.

XX 01-JUN-2000; 2000MO-US14973.

XX 11-JUN-1999; 99US-0138630.

XX (HUMA-) HUMAN GENOME SCI INC.

PA (ROSE/) ROSEN C A.

PI Rosen CA, Ruben SM, Komatsoulis GA;

XX WPI: 2001-071258/08.

DR N-PSDB; AAF64221.

XX Nucleic acid molecules encoding human secreted proteins, used in

PT preventing, treating or ameliorating a disorder, e.g. Alzheimer's and

XX Parkinson's diseases and cancers -

XX Disclosure: Page 94; 542pp; English.

XX Human secreted proteins AAB75506 - AAB75554 are encoded by polynucleotide  
CC sequences AAF64176 - AAF64224. The specification includes amino acid  
CC sequences AAB75555 - AAB75606 which represent fragments of the human  
CC secreted proteins, and protein sequences with which they share homology.  
CC The proteins and polynucleotides, their agonists and antagonists have  
CC activities dependent on the tissues and cells in which they are  
CC expressed, examples of these activities include, immunosuppressive;  
CC antiarthritic; antirheumatic; antiproliferative; cytostatic; cardiant;  
CC vasotropic; cerebroprotective; neurotropic; neuroprotective; antibacterial;  
CC virucide; fungicide; ophthalmological; and vulnery. The proteins  
CC polynucleotides, agonists and antagonists can be used to treat or detect  
CC or diagnose various diseases and disorders including, autoimmune  
CC diseases e.g. rheumatoid arthritis, hyperproliferative disorders  
CC e.g. neoplasms of the breast or liver, cardiovascular disorders  
CC e.g. cardiac arrest, cerebrovascular disorders e.g. cerebral ischaemia,  
CC angiogenesis, nervous system disorders e.g. Alzheimer's disease,  
CC infections caused by bacteria, viruses and fungi and ocular disorders

CC e.g. corneal infection. The polypeptides can also be used to aid wound  
CC healing and epithelial cell proliferation, to prevent skin ageing due to  
CC sunburn, to maintain organs before transplantation, for supporting cell  
CC culture of primary tissues, to regenerate tissues and in chemotaxis. The  
CC polypeptides can also be used as a food additive or preservative to  
CC increase or decrease storage capabilities. Included in the invention are  
CC polynucleotide sequences AAF64167 - AAF64175 and peptide AAB75505 which  
CC are used in the isolation, identification and characterisation of the  
CC proteins of the invention.

XX Sequence 203 AA:

Query Match 30.5%; Score 62; DB 22; Length 203;  
Best Local Similarity 42.4%; Pred. No. 0.22;  
Matches 14; Conservative 5; Mismatches 14; Indels 0; Gaps 0;

OY 1 IVHGDRKRTAIGLAIAIGLGIFTLCQAVEYVE 33  
Db 93 Ilenmrnglqallllllllylllgaseyfe 125

RESULT 5  
AAV35610  
ID AAV35610 standard; Protein: 455 AA.

AC AAV35610;

DT 13-SEP-1999 (first entry)

DE Chlamydia pneumoniae lipoprotein sequence.

XX Respiratory disease; pneumonia; bronchitis; heart disease; sarcoidosis;  
KM sinusitis; purulent otitis media; erythema nodosum; pharyngitis;  
KM vaccine; neutralising epitope.

XX Chlamydia pneumoniae.

OS NC0927105-A2.

PN 03-JUN-1999.

XX 20-NOV-1998; 98MO-1B01890.

XX 04-NOV-1998; 98US-0107078.

XX 21-NOV-1997; 97FR-0014673.

XX (GEST ) GENSET.

PA Griffiths R;

XX WPI: 1999-357842/30.

XX Genome sequence of Chlamydia pneumoniae

PT Page 1338-1339; Disclosure: 1912pp; English.

XX AAV34584-Y35879 represent the proteins encoded by all the open reading  
CC frames in the complete genome (see AAV34584) of Chlamydia pneumoniae.  
CC C. pneumoniae causes respiratory disease such as pneumonia and  
CC bronchitis and is thought to be a contributing factor in heart  
CC disease, sarcoidosis, sinusitis, purulent otitis media, erythema  
CC nodosum or pharyngitis. The polypeptides encoded by the open reading  
CC frames of the C. pneumoniae genome (see AAV34584-Y35879) can be used in  
CC immunogenic compositions as vaccines. Vectors containing C. pneumoniae  
CC nucleotide sequences can also be used as immunogenic compositions,  
CC especially where the vector directs the expression of a neutralising  
CC epitope of C. pneumoniae.

XX Sequence 455 AA:

Query Match 30.5%; Score 62; DB 20; Length 455;

Best Local Similarity 50.0%; Pred. No. 0.56;  
Matches 9; Conservative 6; Mismatches 3; Indels 0; Gaps 0;  
QY 21 WFTFLCOAYEYEVTE 38  
Db 35 WYLCQSYEHKKLVFPE 52

RESULT 6  
AAB29909  
ID AAB29909 standard; Protein; 133 AA.  
AC AAB29909;  
DT 09-FEB-2001 (first entry)  
DE Human secreted protein BLAST search protein SEQ ID NO: 167.

XX Cytostatic; immunosuppressive; nootropic; neuroprotective; antiviral;  
XX anti-allergic; hepatotropic; antidiabetic; antinflammatory; antitumor;  
XX vulnery; anticonvulsant; antibacterial; antifungal; antiparasitic;  
XX cardiac; gene therapy; cancer; immune disorder; cardiovascular disorder;  
XX neurological disease; infection; human; secreted protein.

OS Homo sapiens.  
XX  
XX WO200061779-A1.  
XX  
XX 19-OCT-2000.

XX 06-APR-2000; 2000WO-US09068.

XX 09-APR-1999; 99US-0128699.  
XX 20-JAN-2000; 2000US-0177050.

XX (HUMA-) HUMAN GENOME SCI INC.

PI Rosen CA, Ruben SM, Komatsoulis G;

DR WPI: 2000-647424/62.

XX Isolated nucleic acid molecule encoding a human secreted protein is  
XX used in preventing, treating or ameliorating a medical condition -  
XX

PS Disclosure: Page 487; 495pp; English.

XX The invention relates to the isolation of genes AACC3410-C63458 encoding  
XX the human secreted proteins AAB29802-B29850. This sequence represents a  
XX peptide fragment homologous to the protein encoded by the gene given  
XX in the descriptor line. The sequence is a search result from a BLASTX  
XX homology search. The genes and proteins are useful for preventing,  
XX ameliorating or treating medical conditions, e.g. by protein or gene  
XX therapy. The genes are isolated from a range of human tissues disclosed  
XX in the specification. The nucleic acids, proteins, antibodies and  
XX (ant)agonists are useful in the diagnosis, treatment and prevention of:  
XX (a) cancer, e.g. breast and ovarian cancer, and other cancers of the  
XX adrenal gland, bone, bone marrow, breast, gastrointestinal tract, liver,  
XX lung, or urogenital; (b) immune disorders e.g. Addison's disease,  
XX allergies, autoimmune haemolytic anaemia, autoimmune thyroiditis,  
XX diabetes mellitus, Crohn's disease, multiple sclerosis, rheumatoid  
XX arthritis and ulcerative colitis; (c) cardiovascular disorders such  
XX as myocardial ischaemias; (d) wound healing; (e) neurological diseases  
XX e.g. cerebral anoxia and epilepsy; and (f) infectious diseases such  
XX as viral, bacterial, fungal and parasitic infections.

XX Sequence 133 AA;

Query Match 29.6%; Score 60; DB 21; Length 133;  
Best Local Similarity 42.4%; Pred. No. 0.27;  
Matches 14; Conservative 5; Mismatches 14; Indels 0; Gaps 0;

OY 1 IVHGDRKTAIGLAIGWIFLCOAYEY 33

Db 23 IImennrnmigalllillllyflllgaseyfe 55

RESULT 7  
AAB36196  
ID AAB36196 standard; Protein; 172 AA.  
AC AAB36196;  
DT 02-MAR-2001 (first entry)  
DE Cowdria ruminatum 4hworfl polypeptide.

XX Cowdria ruminatum; MAP1; major antigenic protein 1; antirickettsial;  
XX vaccine; gene therapy; Rickettsia; Cowdria; Anaplasma; map2; 1hworfl3;  
XX 4hworfl; 18hworfl; 3gforfl3.

OS Cowdria ruminatum.  
XX  
XX WO200065063-A2.  
XX  
XX 02-NOV-2000.

XX 21-APR-2000; 2000WO-US10886.

XX 22-APR-1999; 99US-0130725.

XX (UYFL) UNIV FLORIDA.

XX Barbet AF, Bowie MV, Ganta RR, Burridge MJ, Mahan SM, McGuire TC;  
XX Rurangirwa FR, Moreland AL, Simbi BH, Whitmore WW, Alleman AR;

XX WPI: 2000-679675/66.  
XX N-PSDB; AAC68713.

XX New polynucleotides useful as DNA vaccines for conferring immunity to  
XX rickettsial infection e.g. heartwater caused by Cowdria ruminatum,  
XX comprises major antigenic protein 1 or 2 gene of rickettsial pathogens  
XX

PS Claim 3; Page 58-59; 63pp; English.

XX The present sequence is given in a specification relating to nucleic  
XX acid vaccines which may be used to protect animals or humans against  
XX rickettsial diseases caused by a organisms of Rickettsia sp.,  
XX Ehrlichia sp., Anaplasma sp. or Cowdria sp. The vaccine elicits an  
XX immune response protective against the rickettsial pathogen. The  
XX vaccine comprises the major antigenic protein 1 (MAP1) or major  
XX antigenic protein 2 (MAP2) of rickettsial pathogens. The nucleic acid  
XX enhancer-promoter. Cowdria ruminatum genes designated map 2, 1hworfl3,  
XX 4hworfl, 18hworfl and 3gforfl3 may be used in therapeutic and diagnostic  
XX applications. The polypeptides are useful for detecting antibodies  
XX associated with infection by a rickettsial pathogen whilst the  
XX polynucleotides may be used to detect the presence of rickettsial  
XX nucleic acids.

XX Sequence 172 AA;

Query Match 29.6%; Score 60; DB 21; Length 172;  
Best Local Similarity 42.4%; Pred. No. 0.36;  
Matches 14; Conservative 4; Mismatches 13; Indels 2; Gaps 1;

OY 1 IVHGDRK--KTATGALIAIGWIFLCOAYEY 31  
Db 63 ihwdrkyntnigkvtlgnclhaceidgy 95

RESULT 8  
AAB75603  
ID AAB75603 standard; Protein; 203 AA.



OY 3 HDRKK-----TAIGLAIAIGLWITFLC 26  
 DB 27 hgdskeslvsfamghvstvtelwtgnfnvslgtataglgwlltffc 74  
 RESULT 10  
 ID AAY41758 standard; Protein; 747 AA.  
 AC AAY41758;  
 DT 07-DEC-1999 (first entry)  
 DE Human PRO1012 protein sequence.  
 KW Human; PRO; EST; expressed sequence tag; PCR primer; hybridisation;  
 KW probe; blood coagulation disorder; cancer; cellular adhesion disorder;  
 KW secreted protein; transmembrane protein.  
 OS Homo sapiens.  
 PN MO9946281-A2.  
 FD 16-SEP-1999.  
 PF 08-MAR-1999; 99WO-US05028.  
 PR 10-MAR-1998; 98US-0077450.  
 PR 11-MAR-1998; 98US-0077632.  
 PR 11-MAR-1998; 98US-0077641.  
 PR 12-MAR-1998; 98US-0077649.  
 PR 13-MAR-1998; 98US-0077791.  
 PR 17-MAR-1998; 98US-0040220.  
 PR 20-MAR-1998; 98US-0078886.  
 PR 20-MAR-1998; 98US-0078910.  
 PR 20-MAR-1998; 98US-0078936.  
 PR 25-MAR-1998; 98US-0078939.  
 PR 26-MAR-1998; 98US-0079294.  
 PR 27-MAR-1998; 98US-0079656.  
 PR 27-MAR-1998; 98US-0079664.  
 PR 27-MAR-1998; 98US-0079689.  
 PR 30-MAR-1998; 98US-0079786.  
 PR 30-MAR-1998; 98US-0079920.  
 PR 31-MAR-1998; 98US-0080105.  
 PR 31-MAR-1998; 98US-0080107.  
 PR 31-MAR-1998; 98US-0080165.  
 PR 31-MAR-1998; 98US-0080194.  
 PR 01-APR-1998; 98US-0080327.  
 PR 01-APR-1998; 98US-0080328.  
 PR 01-APR-1998; 98US-0080333.  
 PR 08-APR-1998; 98US-0080334.  
 PR 08-APR-1998; 98US-0081049.  
 PR 08-APR-1998; 98US-0081070.  
 PR 09-APR-1998; 98US-0081071.  
 PR 09-APR-1998; 98US-0081195.  
 PR 09-APR-1998; 98US-0081203.  
 PR 15-APR-1998; 98US-0081229.  
 PR 15-APR-1998; 98US-0081817.  
 PR 15-APR-1998; 98US-0081838.  
 PR 15-APR-1998; 98US-0081952.  
 PR 21-APR-1998; 98US-0081953.  
 PR 21-APR-1998; 98US-0082568.  
 PR 22-APR-1998; 98US-0082700.  
 PR 22-APR-1998; 98US-0082704.  
 PR 22-APR-1998; 98US-0082804.  
 PR 23-APR-1998; 98US-0082767.  
 PR 23-APR-1998; 98US-0082796.

PR 27-APR-1998; 98US-0083336.  
 PR 28-APR-1998; 98US-0083322.  
 PR 29-APR-1998; 98US-0083392.  
 PR 29-APR-1998; 98US-0083495.  
 PR 29-APR-1998; 98US-0083496.  
 PR 29-APR-1998; 98US-0083499.  
 PR 29-APR-1998; 98US-0083500.  
 PR 29-APR-1998; 98US-0083545.  
 PR 29-APR-1998; 98US-0083554.  
 PR 29-APR-1998; 98US-0083558.  
 PR 30-APR-1998; 98US-0083559.  
 PR 05-MAY-1998; 98US-0083742.  
 PR 06-MAY-1998; 98US-0084366.  
 PR 06-MAY-1998; 98US-0084414.  
 PR 07-MAY-1998; 98US-0084441.  
 PR 07-MAY-1998; 98US-0084598.  
 PR 07-MAY-1998; 98US-0084600.  
 PR 07-MAY-1998; 98US-0084627.  
 PR 07-MAY-1998; 98US-0084637.  
 PR 07-MAY-1998; 98US-0084639.  
 PR 07-MAY-1998; 98US-0084640.  
 PR 13-MAY-1998; 98US-0084643.  
 PR 13-MAY-1998; 98US-0085323.  
 PR 15-MAY-1998; 98US-0085338.  
 PR 15-MAY-1998; 98US-0085373.  
 PR 15-MAY-1998; 98US-0085579.  
 PR 15-MAY-1998; 98US-0085580.  
 PR 15-MAY-1998; 98US-0085582.  
 PR 15-MAY-1998; 98US-0085589.  
 PR 15-MAY-1998; 98US-0085697.  
 PR 15-MAY-1998; 98US-0085700.  
 PR 18-MAY-1998; 98US-0085704.  
 PR 22-MAY-1998; 98US-0086023.  
 PR 22-MAY-1998; 98US-0086029.  
 PR 22-MAY-1998; 98US-0086414.  
 PR 22-MAY-1998; 98US-0086417.  
 PR 28-MAY-1998; 98US-0086486.  
 PR 28-MAY-1998; 98US-0087098.  
 PR 28-MAY-1998; 98US-0087106.  
 PR 30-JUL-1998; 98US-0087208.  
 PR 11-SEP-1998; 98US-0100038.  
 PA (GETH ) GENENTECH INC.  
 PI Wood WI, Goddard A, Gurney A, Yuan J, Baker KP, Chen J;  
 DR WPI; 1999-551358/46.  
 DR N-PSDB; AA234275.  
 XX  
 PT New secreted and transmembrane polypeptides and their polynucleotides,  
 PT useful for treating blood coagulation disorders, cancers and cellular  
 PT adhesion disorders -  
 PS  
 PS Claim 12; Fig 190; 530pp; English.  
 XX  
 CC The present invention describes secreted and transmembrane polypeptides  
 CC and their polynucleotides. The nucleotide sequences are useful as  
 CC sources of probes, primers, for chromosome mapping, and for generation  
 CC of antisense sequences. They can also be used to treat a variety of diseases  
 CC disorders, depending on their function. Diseases that may be treated  
 CC include blood coagulation disorders, cancers and cellular adhesion  
 CC disorders. They may also be used to raise antibodies. AA233891 to  
 CC AA234338, and AAY41774 represent polynucleotide and  
 CC polypeptide sequence given in the exemplification of the present  
 CC invention.  
 XX  
 SQ Sequence 747 AA;

Query Match

28.6%; Score 58; DB 20; Length 747;

Best Local Similarity 27.1%; Pred. No. 4;  
Matches 13; Conservative 3; Mismatches 8; Indels 24; Gaps 1;

3 HGDRKK-----TAIGLAIAIGLGIPTLC 26  
| | | | | : | | | | |  
217 hgdrskesjvsfamghvstvtelwtgnfnvsigtatfaagigwiltfc 264

RESULT 11  
AAB44314  
ID AAB44314 standard; Protein; 747 AA.  
AC AAB44314;  
XX  
XX 08-FEB-2001 (first entry)  
DE Human PRO1012 (UNQ495) protein sequence SEQ ID NO:459.  
XX  
XX Human; secreted protein; transmembrane protein; PRO; EST; cytosolic;  
KW expressed sequence tag; detection; cancer.  
XX  
XX Homo sapiens.  
OS  
PN MO200053756-A2.  
XX  
PD 14-SEP-2000.  
XX  
PF 18-FEB-2000; 2000WO-US04311.  
XX  
XX 08-MAR-1999; 99WO-US05028.  
PR 12-MAR-1999; 99US-0123957.  
PR 29-MAR-1999; 99US-0126773.  
PR 21-APR-1999; 99US-0130232.  
PR 28-APR-1999; 99US-0131445.  
PR 14-MAY-1999; 99US-0134287.  
PR 23-JUN-1999; 99US-0141037.  
PR 26-JUL-1999; 99US-0145698.  
PR 29-OCT-1999; 99US-0162506.  
PR 30-NOV-1999; 99WO-US28313.  
PR 02-DEC-1999; 99WO-US28551.  
PR 02-DEC-1999; 99WO-US28565.  
PR 16-DEC-1999; 99WO-US30095.  
PR 30-DEC-1999; 99WO-US31243.  
PR 30-DEC-1999; 99WO-US31274.  
PR 05-JAN-2000; 2000WO-US00219.  
PR 06-JAN-2000; 2000WO-US00277.  
PR 06-JAN-2000; 2000WO-US00376.  
XX  
XX (GETH ) GENENTECH INC.  
XX  
PA Ashkenazi AJ, Baker KP, Botstein D, Desnoyers L, Eaton DL;  
PI Ferrara N, Filvaroff E, Fong S, Gao W, Gerber H, Gerritsen ME;  
PI Goddard A, Godowski PJ, Grimaldi CJ, Gurney AL, Hillan KJ;  
PI Kijavini IJ, Kuo SS, Nepier MA, Pan J, Paoni NF, Roy MA;  
PI Shelton DL, Stewart TA, Tumas D, Williams PM, Wood WI;  
XX  
XX WPI: 2000-611443/58.  
DR N-PSDB: AAC78573.  
XX  
XX Novel PRO polypeptides and polynucleotides used in detection methods,  
PT to target bioactive molecules to specific cells, and to modulate  
P7 cellular activities -  
XX  
XX Claim 12; Fig 190; 636pp; English.

CC The polypeptide pairs provide specific targeting of bioactive molecules  
CC to cells. AAC78600 to AAC78987 represent PCR primers and probes used in  
CC the isolation of the PRO polynucleotide sequences.  
XX  
XX Sequence 747 AA;  
SO

Query Match 28.6%; Score 58; DB 21; Length 747;  
Best Local Similarity 27.1%; Pred. No. 4;  
Matches 13; Conservative 3; Mismatches 8; Indels 24; Gaps 1;

3 HGDRKK-----TAIGLAIAIGLGIPTLC 26  
| | | | | : | | | | |  
217 hgdrskesjvsfamghvstvtelwtgnfnvsigtatfaagigwiltfc 264

RESULT 12  
AAU12384  
ID AAU12384 standard; Protein; 747 AA.  
AC AAU12384;  
XX  
XX 24-OCT-2001 (first entry)  
DE Human PRO1012 polypeptide sequence.  
XX  
XX Human secretory and transmembrane; PRO; mammalian; cancer; lung;  
KW breast; prostate; cervical; tumour necrosis factor-alpha; TNF-alpha;  
KW cartilage; ear; proliferation; glucose; free fatty acid; skeletal muscle;  
KW adipocyte; A-peptide; factor VIIa; gene therapy.  
XX  
XX Homo sapiens.  
OS  
PN WO200140466-A2.  
XX  
PD 07-JUN-2001.  
XX  
PF 01-DEC-2000; 2000WO-US32678.  
XX  
XX 01-DEC-1999; 99WO-US28301.  
PR 01-DEC-1999; 99WO-US28634.  
PR 02-DEC-1999; 99WO-US28551.  
PR 02-DEC-1999; 99WO-US28564.  
PR 02-DEC-1999; 99WO-US28565.  
PR 09-DEC-1999; 99US-0170262.  
PR 16-DEC-1999; 99WO-US30095.  
PR 20-DEC-1999; 99WO-US30911.  
PR 20-DEC-1999; 99WO-US30999.  
PR 30-DEC-1999; 99WO-US31243.  
PR 06-JAN-2000; 2000WO-US00277.  
PR 06-JAN-2000; 2000WO-US00376.  
PR 11-FEB-2000; 2000WO-US03565.  
PR 18-FEB-2000; 2000WO-US04341.  
PR 18-FEB-2000; 2000WO-US04342.  
PR 22-FEB-2000; 2000WO-US04414.  
PR 24-FEB-2000; 2000WO-US04914.  
PR 24-FEB-2000; 2000WO-US05004.  
PR 01-MAR-2000; 2000WO-US05601.  
PR 20-MAR-2000; 2000WO-US07377.  
PR 21-MAR-2000; 2000WO-US07532.  
PR 30-MAR-2000; 2000WO-US08439.  
PR 17-MAY-2000; 2000WO-US13705.  
PR 22-MAY-2000; 2000WO-US14042.  
PR 30-MAY-2000; 2000WO-US14941.  
PR 02-JUN-2000; 2000WO-US15264.  
PR 10-NOV-2000; 2000WO-US30873.  
XX  
XX (GETH ) GENENTECH INC.  
XX  
PA Baker KP, Beresini M, Deforge L, Desnoyers L, Filvaroff E, Gao W;  
PI Gerritsen ME, Goddard A, Godowski PJ, Gurney AL, Sherwood S;  
PI Smith V, Stewart TA, Tumas D, Watanabe CK, Wood WI, Zhang Z;

DR WPI: 2001-408281/43.  
DR N-PSDB: AAS21456.

XX Isolated, secretory and transmembrane PRO polypeptide used to detect  
PT other PRO polypeptides, link bioactive molecules to cells expressing  
PT PRO polypeptides, and detect the presence of mammalian tumours e.g.  
PT lung, breast, prostate, cervical

XX Claim 12; Fig 426; 813pp; English.

XX AAU12172-AAU12446 represent novel human secretory and transmembrane  
CC PRO polypeptides. The PRO polypeptides are useful to detect other  
CC PRO polypeptides, to link bioactive molecules to cells expressing  
CC PRO polypeptides, to modulate biological activities of cells expressing  
CC breast, prostate, rectal, cervical or liver tumours by comparing PRO  
CC polypeptide expression in a cell sample to that in a control sample.  
CC Some of the 275 sequences are also useful to stimulate the release of  
CC tumor necrosis factor-alpha (TNF-alpha) from human blood, the  
CC proliferation or differentiation of chondrocytes, the proliferation or  
CC gene expression in pericyte cells, the release of proteoglycans from  
CC cartilage, the proliferation of inner ear utricular supporting cells or  
CC of T-lymphocytes, or the release of a cytokine from peripheral blood  
CC monocytes (PBMCs), or the proliferation of endothelial cells. Some of  
CC the PRO polypeptides may modulate glucose or free fatty acid uptake by  
CC skeletal muscle cells or by adipocytes; or inhibit binding of A-peptide  
CC to factor VIIA. The PRO polypeptides can be used in assays to identify  
CC molecules involved in binding interactions. The polynucleotides encoding  
CC PRO polypeptides can be used to generate probes, antisense RNA/DNA,  
CC transgenic or knock out animals and can be used in gene therapy.

XX Sequence 747 AA;

Query Match 28.6%; Score 58; DB 22; Length 747;  
Best Local Similarity 27.1%; Pred. No. 4;  
Matches 13; Conservative 3; Mismatches 8; Indels 24; Gaps 1;

QY 3 HGDRKK-----TAIGLAIAIGLWIFPLC 26  
Db 217 hgdrskeslvsfamghvstvtelwtgntfnsiqtafaaglwltfc 264

RESULT 13

AA92743  
ID AAB92743 standard; Protein; 768 AA.

AC AAB92743;

DT 26-JUN-2001 (first entry)

DE Human protein sequence SEQ ID NO:11190.

KW Human; primer: detection; diagnosis; antisense therapy; gene therapy.

OS Homo sapiens.

PN EPI074617-A2.

XX 07-FEB-2001.

PF 28-JUL-2000; 2000EP-0116126.

PR 29-JUL-1999; 99JP-0248036.

PR 27-AUG-1999; 99JP-0300253.

PR 11-JAN-2000; 2000JP-0118776.

PR 02-MAY-2000; 2000JP-0183767.

PR 09-JUN-2000; 2000JP-0241899.

PA (HELI-) HELIX RES INST.

PI Ota T, Isogai T, Nishikawa T, Hayashi K, Saito K, Yamamoto J;  
PI Ishii S, Sugiyama T, Wakamatsu A, Nagai K, Otsuki T;

XX WPI: 2001-318749/34.

XX Primer sets for synthesizing polynucleotides, particularly the 5602  
PT full-length cDNAs defined in the specification, and for the detection  
PT and/or diagnosis of the abnormality of the proteins encoded by the  
PT full-length cDNAs -

XX Claim 8; SEQ ID 11190; 2537pp + CD ROM; English.

XX The present invention describes primer sets for synthesizing 5602  
CC full-length cDNAs defined in the specification, where a primer set  
CC comprises: (a) an oligo-dT primer and an oligonucleotide complementary  
CC to the complementary strand of a polynucleotide which comprises one of  
CC the 5602 nucleotide sequences defined in the specification, where the  
CC of an oligonucleotide comprises at least 15 nucleotides; or (b) a combination  
CC of an oligonucleotide comprising a sequence complementary to the  
CC complementary strand of a polynucleotide comprising a sequence complementary to a  
CC polynucleotide which comprises a 3'-end sequence, where the  
CC oligonucleotide comprises at least 15 nucleotides and the combination of  
CC the 5'-end sequence/3'-end sequence is selected from those defined in  
CC the specification. The primer sets can be used in antisense therapy and  
CC in gene therapy. The primers are useful for synthesizing polynucleotides,  
CC particularly full-length cDNAs. The primers are also useful for the  
CC detection and/or diagnosis of the abnormality of the proteins encoded by  
CC the full-length cDNAs. The primers allow obtaining of the full-length  
CC cDNAs easily without any specialised methods. AAH03166 to AAH13628 and  
CC AAH13633 to AAH18742 represent human cDNA sequences; AAB92446 to  
CC AAB95893 represent human amino acid sequences; and AAH13629 to AAH13632  
CC represent oligonucleotides, all of which are used in the exemplification  
CC of the present invention.

XX Sequence 768 AA;

Query Match 28.6%; Score 58; DB 22; Length 768;  
Best Local Similarity 27.1%; Pred. No. 4.1;  
Matches 13; Conservative 3; Mismatches 8; Indels 24; Gaps 1;

QY 3 HGDRKK-----TAIGLAIAIGLWIFPLC 26  
Db 192 hgdrskeslvsfamghvstvtelwtgntfnsiqtafaaglwltfc 239

RESULT 14

AA93204  
ID AAB93204 standard; Protein; 793 AA.

AC AAB93204;

DT 26-JUN-2001 (first entry)

DE Human protein sequence SEQ ID NO:12172.

KW Human; primer: detection; diagnosis; antisense therapy; gene therapy.

OS Homo sapiens.

PN EPI074617-A2.

XX 07-FEB-2001.

PF 28-JUL-2000; 2000EP-0116126.

PR 29-JUL-1999; 99JP-0248036.

PR 27-AUG-1999; 99JP-0300253.

PR 11-JAN-2000; 2000JP-0118776.

PR 02-MAY-2000; 2000JP-0183767.

PR 09-JUN-2000; 2000JP-0241899.

PA (HELI-) HELIX RES INST.

PI Ota T, Isogai T, Nishikawa T, Hayashi K, Saito K, Yamamoto J;  
 PI Ishii S, Sugiyama T, Wakamatsu A, Nagai K, Otsuki T;  
 DR WPI: 2001-318749/34.

PR Primer sets for synthesizing polynucleotides, particularly the 5602  
 PT full-length cDNAs defined in the specification, and for the detection  
 PT and/or diagnosis of the abnormality of the proteins encoded by the  
 PT full-length cDNAs -

PS Claim 8; SEQ ID 12172; 2537pp + CD ROM; English.

XX The present invention describes primer sets for synthesizing 5602  
 CC full-length cDNAs defined in the specification. Where a primer set  
 CC comprises: (a) an oligo-dT primer and an oligonucleotide complementary  
 CC to the complementary strand of a polynucleotide which comprises one of  
 CC the 5602 nucleotide sequences defined in the specification, where the  
 CC oligonucleotide comprises at least 15 nucleotides; or (b) a combination  
 CC of an oligonucleotide comprising a sequence complementary to the  
 CC complementary strand of a polynucleotide which comprises a 5'-end  
 CC sequence and an oligonucleotide comprising a sequence complementary to a  
 CC polynucleotide which comprises a 3'-end sequence, where the  
 CC oligonucleotide comprises at least 15 nucleotides and the combination of  
 CC the 5'-end sequence/3'-end sequence is selected from those defined in  
 CC the specification. The primer sets can be used in antisense therapy and  
 CC in gene therapy. The primers are useful for synthesizing polynucleotides,  
 CC particularly full-length cDNAs. The primers are also useful for the  
 CC detection and/or diagnosis of the abnormality of the proteins encoded by  
 CC the full-length cDNAs. The primers allow obtaining of the full-length  
 CC cDNAs easily without any specialised methods. AAH03166 to AAH13628 and  
 CC AAH13633 to AAH18742 represent human cDNA sequences; AAB92446 to  
 CC AAB95893 represent human amino acid sequences; and AAH13629 to AAH13632  
 CC represent oligonucleotides, all of which are used in the exemplification  
 CC of the present invention.

CC Sequence 793 AA;

Query Match 28.6%; Score 58; DB 22; Length 793;

Best Local Similarity 27.1%; Pred. No. 4.3;

Matches 13; Conservative 3; Mismatches 8; Indels 24; Gaps 1;

OY 3 HGDRKK-----TAIGLAIAIGLWIFLTC 26  
 |||||  
 Db 217 hgdskeslvsfamghvrstvtelwtgnfvnsiqtafaagigwiltfc 264

RESULT 15

AAB95309

XX ID AAB95309 standard; Protein; 793 AA.

XX AC AAB95309;

XX DT 26-JUN-2001 (first entry)

XX DE Human protein sequence SEQ ID NO:17552.

XX KM Human; primer; detection; diagnosis; antisense therapy; gene therapy.

XX OS Homo sapiens.

XX PN EP1074617-A2.

XX PD 07-FEB-2001.

XX PF 28-JUL-2000; 2000EP-0116126.

XX PR 29-JUL-1999; 99JP-0248036.

XX PR 27-AUG-1999; 99JP-0300253.

XX PR 11-JAN-2000; 2000JP-0118776.

XX PR 02-MAY-2000; 2000JP-0183767.

XX PR 09-JUN-2000; 2000JP-0241899.

PA (HELI-) HELIX RES INST.

XX Ota T, Isogai T, Nishikawa T, Hayashi K, Saito K, Yamamoto J;  
 PI Ishii S, Sugiyama T, Wakamatsu A, Nagai K, Otsuki T;  
 DR WPI: 2001-318749/34.

PR Primer sets for synthesizing polynucleotides, particularly the 5602  
 PT full-length cDNAs defined in the specification, and for the detection  
 PT and/or diagnosis of the abnormality of the proteins encoded by the  
 PT full-length cDNAs -

PS Claim 8; SEQ ID 17552; 2537pp + CD ROM; English.

XX The present invention describes primer sets for synthesizing 5602  
 CC full-length cDNAs defined in the specification. Where a primer set  
 CC comprises: (a) an oligo-dT primer and an oligonucleotide complementary  
 CC to the complementary strand of a polynucleotide which comprises one of  
 CC the 5602 nucleotide sequences defined in the specification, where the  
 CC oligonucleotide comprises at least 15 nucleotides; or (b) a combination  
 CC of an oligonucleotide comprising a sequence complementary to the  
 CC complementary strand of a polynucleotide which comprises a 5'-end  
 CC sequence and an oligonucleotide comprising a sequence complementary to a  
 CC polynucleotide which comprises a 3'-end sequence, where the  
 CC oligonucleotide comprises at least 15 nucleotides and the combination of  
 CC the 5'-end sequence/3'-end sequence is selected from those defined in  
 CC the specification. The primer sets can be used in antisense therapy and  
 CC in gene therapy. The primers are useful for synthesizing polynucleotides,  
 CC particularly full-length cDNAs. The primers are also useful for the  
 CC detection and/or diagnosis of the abnormality of the proteins encoded by  
 CC the full-length cDNAs. The primers allow obtaining of the full-length  
 CC cDNAs easily without any specialised methods. AAH03166 to AAH13628 and  
 CC AAH13633 to AAH18742 represent human cDNA sequences; AAB92446 to  
 CC AAB95893 represent human amino acid sequences; and AAH13629 to AAH13632  
 CC represent oligonucleotides, all of which are used in the exemplification  
 CC of the present invention.

XX Sequence 793 AA;

Query Match 28.6%; Score 58; DB 22; Length 793;

Best Local Similarity 27.1%; Pred. No. 4.3;

Matches 13; Conservative 3; Mismatches 8; Indels 24; Gaps 1;

OY 3 HGDRKK-----TAIGLAIAIGLWIFLTC 26  
 |||||  
 Db 217 hgdskeslvsfamghvrstvtelwtgnfvnsiqtafaagigwiltfc 264

Search completed: January 10, 2002, 02:01:16  
 Job time: 289 sec







Db 151 LMENNRNOMIALITLILGLYFTLLQASEYFE 183

## RESULT 2

PCT-US91-09422-19  
; Sequence 19, Application PC/TUS9109422

## GENERAL INFORMATION:

APPLICANT: MULVIHILL, Eileen R.  
APPLICANT: Hagen, Frederick S.  
APPLICANT: Houamed, Khaled M.  
APPLICANT: Almers, Wolfhard  
TITLE OF INVENTION: G PROTEIN-COUPLED GLUTAMATE RECEPTORS  
NUMBER OF SEQUENCES: 33  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: Townsend and Townsend  
STREET: One Market Plaza, Stewart Street Tower  
CITY: San Francisco  
STATE: California  
COUNTRY: USA  
ZIP: 94105-1492

## COMPUTER READABLE FORM:

MEDIUM TYPE: Floppy disk  
COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: Patentin Release #1.0, Version #1.25  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: PCT/US91/09422  
FILING DATE: 19911212  
CLASSIFICATION: 435PRIOR APPLICATION DATA:  
APPLICATION NUMBER: US 07/672,007  
FILING DATE: 18-MAR-1991APPLICATION NUMBER: US 07/648,481  
FILING DATE: 30-JAN-1991PRIOR APPLICATION DATA:  
APPLICATION NUMBER: US 07/626,806  
FILING DATE: 12-DEC-1990ATTORNEY/AGENT INFORMATION:  
NAME: Parmelee, Steven W.  
REGISTRATION NUMBER: 31,990REFERENCE/DOCKET NUMBER: 13952-6PC  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: (206) 467-9600TELEFAX: (415) 543-5043  
INFORMATION FOR SEQ ID NO: 19:

SEQUENCE CHARACTERISTICS:

LENGTH: 912 amino acids

TYPE: AMINO ACID

TOPOLOGY: linear

MOLECULE TYPE: protein  
PCT-US91-09422-19Query Match 24.6%; Score 50; DB 5; Length 912;  
Best Local Similarity 39.3%; Pred. No. 35;

Matches 11; Conservative 3; Mismatches 10; Indels 4; Gaps 1;

QY 4 GDRKRTAIGLAIGLWFTLLCOAYEY 31

Db 527 GERKRYKGMAC---WHCEPCTGYOY 550

## RESULT 3

US-08-460-269C-2

; Sequence 2, Application US/08460269C  
Patent No. 6197548

## GENERAL INFORMATION:

APPLICANT: CLARE, JEFFREY J.  
ROMANOS, MICHAEL A.  
TITLE OF INVENTION: EXPRESSION OF HETEROLOGOUS PROTEIN IN  
YEAST  
NUMBER OF SEQUENCES: 17

## CORRESPONDENCE ADDRESS:

ADDRESSEE: Millen, White, Zelano & Branigan, P.C.  
STREET: 2200 Clarendon Blvd., Suite 1400  
CITY: ARLINGTON  
STATE: VA  
COUNTRY: USA  
ZIP: 22201

## COMPUTER READABLE FORM:

MEDIUM TYPE: Floppy disk  
COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: Patentin Release #1.0, Version #1.30  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/08/460,269C  
FILING DATE: 02-Jun-1995ATTORNEY/AGENT INFORMATION:  
NAME: Lebovitz, Richard M.

REGISTRATION NUMBER: 37,067

REFERENCE/DOCKET NUMBER: Popov-2

TELECOMMUNICATION INFORMATION:  
TELEPHONE: (703) 243-6333

TELEFAX: (703) 243-6410

INFORMATION FOR SEQ ID NO: 2:

SEQUENCE CHARACTERISTICS:

LENGTH: 910 amino acids

TYPE: amino acid

TOPOLOGY: linear

MOLECULE TYPE: protein

SEQUENCE DESCRIPTION: SEQ ID NO: 2:  
US-08-460-269C-2Query Match 24.1%; Score 49; DB 4; Length 910;  
Best Local Similarity 36.7%; Pred. No. 49;

Matches 11; Conservative 7; Mismatches 10; Indels 2; Gaps 1;

QY 2 VHGDRKRTAIGLAIGLWFTLLCOAYEY 31

Db 863 LRGTRAEIGLGMALGSG--HSLYASYEY 890

## RESULT 4

US-08-460-269C-4

; Sequence 4, Application US/08460269C  
Patent No. 6197548

## GENERAL INFORMATION:

APPLICANT: CLARE, JEFFREY J.  
ROMANOS, MICHAEL A.TITLE OF INVENTION: EXPRESSION OF HETEROLOGOUS PROTEIN IN  
YEAST

NUMBER OF SEQUENCES: 17

CORRESPONDENCE ADDRESS:

ADDRESSEE: Millen, White, Zelano &amp; Branigan, P.C.

STREET: 2200 Clarendon Blvd., Suite 1400

CITY: ARLINGTON

STATE: VA

COUNTRY: USA  
ZIP: 22201

## COMPUTER READABLE FORM:

MEDIUM TYPE: Floppy disk  
COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: Patentin Release #1.0, Version #1.30  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/08/460,269C  
FILING DATE: 02-Jun-1995

ATTORNEY/AGENT INFORMATION:

NAME: Lebovitz, Richard M.

REGISTRATION NUMBER: 37,067

REFERENCE/DOCKET NUMBER: Popov-2

TELECOMMUNICATION INFORMATION:  
TELEPHONE: (703) 243-6333

TELEFAX: (703) 243-6410

INFORMATION FOR SEQ ID NO: 4:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 911 amino acids  
TYPE: amino acid  
TOPOLOGY: linear  
MOLECULE TYPE: protein  
SEQUENCE DESCRIPTION: SEQ ID NO: 4:  
US-08-460-269C-4

Query Match 24.1%; Score 49; DB 4; Length 911;  
Best Local Similarity 36.7%; Pred. No. 49;  
Matches 11; Conservative 7; Mismatches 10; Indels 2; Gaps 1;

QY 2 VHGDRKTAIGLAIGLWIFLCOAYEY 31  
DB 864 LRGTRAEGLGMAALGRG--HSLVASYEY 891

RESULT 5  
US-08-460-269C-6  
Sequence 6, Application US/08460269C  
Patent No. 6197548  
GENERAL INFORMATION:  
APPLICANT: CLARE, JEFFREY J.  
ROMANOS, MICHAEL A.  
TITLE OF INVENTION: EXPRESSION OF HETEROLOGOUS PROTEIN IN YEAST  
NUMBER OF SEQUENCES: 17  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: Millen, White, Zelano & Brannigan, P.C.  
STREET: 2200 Clarendon Blvd., Suite 1400  
CITY: ARLINGTON  
STATE: VA  
COUNTRY: USA  
ZIP: 22201  
COMPUTER READABLE FORM:  
MEDIUM TYPE: Floppy disk  
COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: Patentin Release #1.0, Version #1.30  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/08/460,269C  
FILING DATE: 02-Jun-1995  
ATTORNEY/AGENT INFORMATION:  
NAME: Lebovitz, Richard M.  
REGISTRATION NUMBER: 37,067  
REFERENCE/DOCKET NUMBER: Popov-2  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: (703) 243-6333  
TELEFAX: (703) 243-6410  
INFORMATION FOR SEQ ID NO: 6:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 922 amino acids  
TYPE: amino acid  
TOPOLOGY: linear  
MOLECULE TYPE: protein  
SEQUENCE DESCRIPTION: SEQ ID NO: 6:  
US-08-460-269C-6

Query Match 24.1%; Score 49; DB 4; Length 922;  
Best Local Similarity 36.7%; Pred. No. 50;  
Matches 11; Conservative 7; Mismatches 10; Indels 2; Gaps 1;

QY 2 VHGDRKTAIGLAIGLWIFLCOAYEY 31  
DB 875 LRGTRAEGLGMAALGRG--HSLVASYEY 902

RESULT 6  
US-08-118-441-29  
Sequence 29, Application US/08118441

Patent No. 5578493  
GENERAL INFORMATION:  
APPLICANT: Gilliam, T. Conrad  
TANZI, RUDOLPH E.  
TITLE OF INVENTION: ISOLATION AND USES OF A WILSON'S DISEASE GENE  
NUMBER OF SEQUENCES: 29  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: Cooper & Dunham  
STREET: 30 Rockefeller Plaza  
CITY: New York  
STATE: New York  
COUNTRY: United States of America  
ZIP: 10112  
COMPUTER READABLE FORM:  
MEDIUM TYPE: Floppy disk  
COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: Patentin Release #1.0, Version #1.25  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/08/118,441  
FILING DATE:  
CLASSIFICATION: 435  
ATTORNEY/AGENT INFORMATION:  
NAME: White, John P.  
REGISTRATION NUMBER: 28,678  
REFERENCE/DOCKET NUMBER: 0575/44011  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: (212) 977-9550  
TELEFAX: (212) 664-0525  
INFORMATION FOR SEQ ID NO: 29:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 1110 amino acids  
TYPE: amino acid  
TOPOLOGY: linear  
MOLECULE TYPE: protein  
US-08-118-441-29

Query Match 24.1%; Score 49; DB 1; Length 1110;  
Best Local Similarity 34.8%; Pred. No. 62;  
Matches 8; Conservative 7; Mismatches 8; Indels 0; Gaps 0;

QY 1 IVHGDRKTAIGLAIGLWIF 23  
DB 863 LITGDKRKTAIRAIATOVGINKVF 885

RESULT 7  
US-08-338-579A-29  
Sequence 29, Application US/08338579A  
Patent No. 6068975  
GENERAL INFORMATION:  
APPLICANT: Gilliam, T. Conrad  
TANZI, RUDOLPH E.  
TITLE OF INVENTION: ISOLATION AND USES OF A WILSON'S DISEASE GENE  
NUMBER OF SEQUENCES: 107  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: Cooper & Dunham  
STREET: 1185 Avenue of the Americas  
CITY: New York  
STATE: New York  
COUNTRY: United States of America  
ZIP: 10036  
COMPUTER READABLE FORM:  
MEDIUM TYPE: Floppy disk  
COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: Patentin Release #1.0, Version #1.30  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/08/338,579A



NUMBER OF SEQUENCES: 63  
CORRESPONDENCE ADDRESSES:  
ADDRESSEE: MORRISON & FOERSTER  
STREET: 2000 Pennsylvania Ave. N.W., Ste. 5500  
CITY: Washington  
STATE: D.C.  
COUNTRY: USA  
ZIP: 20006-1812  
COMPUTER READABLE FORM:  
MEDIUM TYPE: Floppy disk  
COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: Patentin Release #1.0, Version #1.30  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/08/476,000  
FILING DATE: 07-JUN-1995  
CLASSIFICATION: 435  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: US 08/390,301  
FILING DATE: 25-JAN-1995  
ATTORNEY/AGENT INFORMATION:  
NAME: ADLER, REID G.  
REGISTRATION NUMBER: 30,988  
REFERENCE/DOCKET NUMBER: 2803-0006.20  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: (202) 887-1500  
TELEFAX: (202) 887-0763  
TELEX: 90-4030  
INFORMATION FOR SEQ ID NO: 63:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 397 amino acids  
TYPE: amino acid  
TOPOLOGY: linear  
MOLECULE TYPE: protein  
US-08-476-000-63

Query Match 22.9%; Score 46.5; DB 1; Length 397;  
Best Local Similarity 38.2%; Pred. No. 42;  
Matches 13; Conservative 5; Mismatches 11; Indels 5; Gaps 2;

OY 6 RKKT-AIGLAIAIGLWITLCQAVEYEIVHT 37  
||| |||::|| |::||  
DB 185 RKKANIAIGISLAI--WLLTLVITPLVVKOT 215

RESULT 11  
US-08-472-840-63  
Sequence 63, Application US/08472840  
Patent No. 5763575  
GENERAL INFORMATION:  
APPLICANT: SUNDELIN, JOHAN  
TITLE OF INVENTION: RECOMBINANT C140 RECEPTOR, ITS AGONISTS  
AND ANTAGONISTS, AND NUCLEIC ACIDS ENCODING THE RECEPTOR  
NUMBER OF SEQUENCES: 63  
CORRESPONDENCE ADDRESSES:  
ADDRESSEE: MORRISON & FOERSTER  
STREET: 2000 Pennsylvania Ave. N.W., Ste. 5500  
CITY: Washington  
STATE: D.C.  
COUNTRY: USA  
ZIP: 20006-1812  
COMPUTER READABLE FORM:  
MEDIUM TYPE: Floppy disk  
COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: Patentin Release #1.0, Version #1.30  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/08/472,840  
FILING DATE:  
CLASSIFICATION: 435  
PRIOR APPLICATION DATA:

APPLICATION NUMBER: US/08/390,301  
FILING DATE: 25-JAN-1995  
ATTORNEY/AGENT INFORMATION:  
NAME: ADLER, REID G.  
REGISTRATION NUMBER: 30,988  
REFERENCE/DOCKET NUMBER: 2803-0006.20  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: (202) 887-1500  
TELEFAX: (202) 887-0763  
TELEX: 90-4030  
INFORMATION FOR SEQ ID NO: 63:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 397 amino acids  
TYPE: amino acid  
TOPOLOGY: linear  
MOLECULE TYPE: protein  
US-08-472-840-63

Query Match 22.9%; Score 46.5; DB 1; Length 397;  
Best Local Similarity 38.2%; Pred. No. 42;  
Matches 13; Conservative 5; Mismatches 11; Indels 5; Gaps 2;

OY 6 RKKT-AIGLAIAIGLWITLCQAVEYEIVHT 37  
||| |||::|| |::||  
DB 185 RKKANIAIGISLAI--WLLTLVITPLVVKOT 215

RESULT 12  
US-08-476-976-63  
Sequence 63, Application US/08476976  
Patent No. 5874400  
GENERAL INFORMATION:  
APPLICANT: SUNDELIN, JOHAN  
TITLE OF INVENTION: RECOMBINANT C140 RECEPTOR, ITS AGONISTS  
AND ANTAGONISTS, AND NUCLEIC ACIDS ENCODING THE RECEPTOR  
NUMBER OF SEQUENCES: 63  
CORRESPONDENCE ADDRESSES:  
ADDRESSEE: MORRISON & FOERSTER  
STREET: 2000 Pennsylvania Ave. N.W., Ste. 5500  
CITY: Washington  
STATE: D.C.  
COUNTRY: USA  
ZIP: 20006-1812  
COMPUTER READABLE FORM:  
MEDIUM TYPE: Floppy disk  
COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: Patentin Release #1.0, Version #1.30  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/08/476,976  
FILING DATE: 07-JUN-1995  
CLASSIFICATION: 536  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: US 08/390,301  
FILING DATE: 25-JAN-1995  
ATTORNEY/AGENT INFORMATION:  
NAME: ADLER, REID G.  
REGISTRATION NUMBER: 30,988  
REFERENCE/DOCKET NUMBER: 2803-0006.20  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: (202) 887-1500  
TELEFAX: (202) 887-0763  
TELEX: 90-4030  
INFORMATION FOR SEQ ID NO: 63:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 397 amino acids  
TYPE: amino acid  
TOPOLOGY: linear  
MOLECULE TYPE: protein  
US-08-476-976-63

Query Match 22.9%; Score 46.5; DB 2; Length 397;  
Best Local Similarity 38.2%; Pred. No. 42;  
Matches 13; Conservative 5; Mismatches 11; Indels 5; Gaps 2;  
QY 6 RKKT--AIGLAIAIGMIFILCOAVEYEYIYHT 37  
||| |||::|| |::| |::|  
Db 185 RKKANIAIGISLAI---WLLTLVTIPLVVKQT 215

RESULT 13  
US-08-474-410-63  
; Sequence 63, Application US/08474410  
; Patent No. 6043212  
; GENERAL INFORMATION:  
; APPLICANT: SUNDELLIN, JOHAN  
; TITLE OF INVENTION: RECOMBINANT C140 RECEPTOR, ITS AGONISTS  
; TITLE OF INVENTION: AND ANTAGONISTS, AND NUCLEIC ACIDS ENCODING THE RECEPTOR  
; NUMBER OF SEQUENCES: 63  
; CORRESPONDENCE ADDRESS:  
; ADDRESSEE: MORRISON & FOERSTER  
; STREET: 2000 Pennsylvania Ave. N.W., Ste. 5500  
; CITY: Washington  
; STATE: D.C.  
; COUNTRY: USA  
; ZIP: 20006-1812  
; COMPUTER READABLE FORM:  
; MEDIUM TYPE: Floppy disk  
; OPERATING SYSTEM: PC-DOS/MS-DOS  
; SOFTWARE: Patentin Release #1.0, Version #1.30  
; CURRENT APPLICATION DATA:  
; APPLICATION NUMBER: US/08/474,410  
; FILING DATE:  
; CLASSIFICATION:  
; PRIOR APPLICATION DATA:  
; APPLICATION NUMBER: US 08/390,301  
; FILING DATE: 25-JAN-1995  
; ATTORNEY/AGENT INFORMATION:  
; NAME: ADLER, REID G.  
; REGISTRATION NUMBER: 30,988  
; REFERENCE/DOCKET NUMBER: 2803-0006.20  
; TELECOMMUNICATION INFORMATION:  
; TELEPHONE: (202) 887-1500  
; TELEFAX: (202) 887-0763  
; TELEX: 90-4030  
; INFORMATION FOR SEQ ID NO: 63:  
; SEQUENCE CHARACTERISTICS:  
; LENGTH: 397 amino acids  
; TYPE: amino acid  
; TOPOLOGY: linear  
; MOLECULE TYPE: protein  
US-08-474-410-63

Query Match 22.9%; Score 46.5; DB 3; Length 397;  
Best Local Similarity 38.2%; Pred. No. 42;  
Matches 13; Conservative 5; Mismatches 11; Indels 5; Gaps 2;  
QY 6 RKKT--AIGLAIAIGMIFILCOAVEYEYIYHT 37  
||| |||::|| |::| |::|  
Db 185 RKKANIAIGISLAI---WLLTLVTIPLVVKQT 215

RESULT 14  
US-08-486-673B-63  
; Sequence 63, Application US/08486673B  
; Patent No. 6297026  
; GENERAL INFORMATION:  
; APPLICANT: Sundellin, Johan  
; APPLICANT: Scarborough, Robert M.  
; TITLE OF INVENTION: Nucleic Acids Encoding the C140 Receptor

FILE REFERENCE: 44481-5006-08-US  
; CURRENT APPLICATION NUMBER: US/08/486,673B  
; CURRENT FILING DATE: 1995-06-07  
; PRIOR APPLICATION NUMBER: US 08/097,938  
; PRIOR FILING DATE: 1993-07-26  
; PRIOR APPLICATION NUMBER: PCT/US94/08536  
; PRIOR FILING DATE: 1994-07-26  
; NUMBER OF SEQ ID NOS: 63  
; SOFTWARE: Patentin Ver. 2.1  
; SEQ ID NO 63  
; LENGTH: 397  
; TYPE: PRT  
; ORGANISM: Homo sapiens  
US-08-486-673B-63

Query Match 22.9%; Score 46.5; DB 4; Length 397;  
Best Local Similarity 38.2%; Pred. No. 42;  
Matches 13; Conservative 5; Mismatches 11; Indels 5; Gaps 2;  
QY 6 RKKT--AIGLAIAIGMIFILCOAVEYEYIYHT 37  
||| |||::|| |::| |::|  
Db 185 RKKANIAIGISLAI---WLLTLVTIPLVVKQT 215

RESULT 15  
US-09-095-758-10  
; Sequence 10, Application US/09095758  
; Patent No. 6184016  
; GENERAL INFORMATION:  
; APPLICANT: PHILIPS, Mark R.  
; TITLE OF INVENTION: HUMAN PRENYLCYSTEINE CARBOXYL  
; TITLE OF INVENTION: METHYLTRANSFERASE, DNA ENCODING SAME, AND A METHOD OF  
; NUMBER OF SEQUENCES: 12  
; CORRESPONDENCE ADDRESS:  
; ADDRESSEE: BROWDY AND NEIMARK, P.L.L.C.  
; STREET: 419 Seventh Street N.W., Ste. 300  
; CITY: Washington  
; STATE: D.C.  
; COUNTRY: U.S.A.  
; ZIP: 20004  
; COMPUTER READABLE FORM:  
; MEDIUM TYPE: Floppy disk  
; OPERATING SYSTEM: IBM PC compatible  
; SOFTWARE: Patentin Release #1.0, Version #1.30  
; CURRENT APPLICATION DATA:  
; APPLICATION NUMBER: US/09/095,758  
; FILING DATE: 11-JUN-1998  
; PRIOR APPLICATION DATA:  
; APPLICATION NUMBER: 60/049,304  
; FILING DATE: 11-JUN-1997  
; ATTORNEY/AGENT INFORMATION:  
; NAME: YON, Allen C.  
; REGISTRATION NUMBER: 37,971  
; REFERENCE/DOCKET NUMBER: PHILIPS-1A  
; TELECOMMUNICATION INFORMATION:  
; TELEPHONE: (202) 628-5197  
; TELEFAX: (202) 737-3528  
; INFORMATION FOR SEQ ID NO: 10:  
; SEQUENCE CHARACTERISTICS:  
; LENGTH: 295 amino acids  
; TYPE: amino acid  
; STRANDEDNESS: single  
; TOPOLOGY: linear  
; MOLECULE TYPE: peptide  
US-09-095-758-10

Query Match 22.7%; Score 46; DB 4; Length 295;  
Best Local Similarity 40.7%; Pred. No. 35;  
Matches 11; Conservative 4; Mismatches 12; Indels 0; Gaps 0;

Thu Jan 10 08:15:45 2002

us-09-712-768-6.raii

Page 7

```

QY      8 KTAIGLAIAIGLWIFTLQAVEYEI 34
      | : | : | : || | : | | |
Db     140 KHSFGYWLASIGWIEFLIENFYPEI 166

```

Search completed: January 10, 2002, 02:01:56  
Job time: 278 sec

1  
1  
1













Align seg 1/1 to: US-09-252-991A-18747 from: 1 to: 230

```

82 CCATCTGCA.....AACGACGCTGTTTGGCGACATC 48
      |||||      ::|||::      |||||
52 ProserArgAlaHisThrGlySerSerArgGlySerAlaProThrSe 68

```

47 GCACCGCGCAGGATCGGATCAGGCAGACCA 17  
 |:::|||||:::||||| |  
 68 rargargAlaAlaCysGlySerMetargPro 78

seq\_name: /cgn2\_6/ptodata/2/paa/PCTUS\_COMB.pep:PCT-US00-06013-142

```
seq_documentation_block:
```

```

; Sequence 142, Application PC/TUS0006013
; GENERAL INFORMATION:

```

APPLICANT: Human Genome Sciences, Inc

```

; TITLE OF INVENTION: 50 Human Secreted Proteins
; FILE REFERENCE: PS508PCT

```

; CURRENT APPLICATION NUMBER: PCT/US00/06013  
 ; CURRENT FILING DATE: 2000-03-09

```
;; EARLIER APPLICATION NUMBER: 60/125,360
; EARLIER FILING DATE: 1999-03-19
```

```

; EARLIER APPLICATION NUMBER: 60/138,626
; EARLIER FILING DATE: 1999-06-11

```

```

; EARLIER APPLICATION NUMBER: 60/168,662
; EARLIER FILING DATE: 1999-12-03

```

```

; NUMBER OF SEQ ID NOS: 170
; SOFTWARE: PatentIn Ver. 2.0

```

```

; SEQ ID NO 142
; LENGTH: 62

```

TYPE: PRT  
ORGANISM: Homo sapiens

NAME/KEY. SITE

LOCATION: (41)  
OTHER INFORMATION

```

; OTHER INFORMATION. add equals sign
;
FEATURE:
NAME/REV:

```

; NAME/KEY: SITE  
; LOCATION: (50)  
; OTHER: INVENTORY

```

;
; OTHER INFORMATION: xaa equals stop translation
;
; FEATURE:
;
;

```

; NAME/KEY: SITE  
; LOCATION: (59)

```
OTHER INFORMATION: xaa equals stop translation
FEATURE:
```

```

; NAME/KEY: SITE
; LOCATION: (60)

```

```

; OTHER INFORMATION: Xaa equals stop translation
PCT-US00-06013-142

```

alignment\_scores: 64.00  
Quality: 64.00

Ratio:	3.200	Gaps:	50 000
Percent similarity:	71 429	Percent identity:	50 000

Income Summary: 11.725 Income Identity: 20.000

US-09-712-768-7 x PCT-US00-06013-142 .

Align seg 1/1 to: PCT-US00-06013-142 from: 1 to: 62

4 TCGATCTTCTGCTGGTCTGCGCTGATCGCATCTGCGGGTGGATGTC 53

```

13 SerThrPheLeuThrIleCysPheIleArgGlnLeuIlePheHisPhetH 29
      ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
      ::::: ::::: ::::: ::::: ::::: ::::: ::::: :::::

```

54 GGCAAACAGCACGTCGGTTTCGAGATGCCGCA 87

```

29 rSerlyshishpheglyphegluaAlaAla 400
      ::::|:::| | | | | | | | | |

```

1

---

GenCore version 4.5  
Copyright (c) 1993 - 2000 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: January 10, 2002, 01:51:17 : Search time 18.96 Seconds  
(without alignments)  
152.671 Million cell updates/sec

Title: us-09-712-768-6  
Sequence: 1 IVHGDRKRTAIGALIAIGLGMIFLCAQAYEYEVHT 38

Scoring table: BLOSUM62  
Gapop 10.0 , Gapext 0.5

Searched: 219241 segs, 76174552 residues  
Total number of hits satisfying chosen parameters: 219241

Minimum DB seq length: 0  
Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%  
Maximum Match 100%  
Listing first 45 summaries

Database :  
1: p1r1:\*  
2: p1r2:\*  
3: p1r3:\*  
4: p1r4:\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

## SUMMARIES

Result No.	Score	Query Match Length	DB ID	Description
1	87	42.9	260 1 OTXL3	cytochrome-c oxida
2	84	41.4	261 2 T10978	cytochrome-c oxida
3	83	40.9	261 2 T11863	cytochrome-c oxida
4	83	40.9	261 2 T11369	cytochrome-c oxida
5	82.5	40.6	266 2 C45164	cytochrome-c oxida
6	81	39.9	261 2 T11283	cytochrome-c oxida
7	81	39.9	274 2 S03807	cytochrome-c oxida
8	80	39.4	261 2 B30396	cytochrome-c oxida
9	80	39.4	261 2 H34012	cytochrome-c oxida
10	80	39.4	261 2 G34012	cytochrome-c oxida
11	80	39.4	261 2 A30396	cytochrome-c oxida
12	80	39.4	261 2 D30396	cytochrome-c oxida
13	80	39.4	261 2 C30396	cytochrome-c oxida
14	80	39.4	261 2 T09953	cytochrome-c oxida
15	79	38.9	261 1 OTB03	cytochrome-c oxida
16	79	38.9	261 1 OTCA3	cytochrome-c oxida
17	79	38.9	261 2 T11486	cytochrome-c oxida
18	79	38.9	261 2 G58888	cytochrome-c oxida
19	77	37.9	261 2 G58850	cytochrome-c oxida
20	77	37.9	261 2 S41826	cytochrome-c oxida
21	77	37.9	261 2 S41841	cytochrome-c oxida
22	76	37.4	261 2 S26157	cytochrome-c oxida
23	76	37.4	261 2 T11499	cytochrome-c oxida
24	76	37.4	261 2 T11253	cytochrome-c oxida
25	76	37.4	261 2 T11408	cytochrome-c oxida
26	75	36.9	261 2 T11081	cytochrome-c oxida
27	75	36.5	261 2 S68134	cytochrome-c oxida
28	74	36.5	261 2 T11512	cytochrome-c oxida
29	74	36.5	261 2 T11512	cytochrome-c oxida

30	74	36.5	261 2 T11447	cytochrome-c oxida
31	74	36.5	261 2 T11540	cytochrome-c oxida
32	74	36.5	261 2 T11146	cytochrome-c oxida
33	74	36.5	261 2 T11434	cytochrome-c oxida
34	74	36.5	262 2 T45556	cytochrome-c oxida
35	72	35.5	261 2 S60277	cytochrome-c oxida
36	72	35.5	261 2 T11826	cytochrome-c oxida
37	72	35.5	261 2 T11343	cytochrome-c oxida
38	71	35.0	261 2 T11173	cytochrome-c oxida
39	71	35.0	261 2 T11306	cytochrome-c oxida
40	71	35.0	262 2 T11883	cytochrome-c oxida
41	71	35.0	270 2 S62760	cytochrome-c oxida
42	70	34.5	261 2 S10193	cytochrome-c oxida
43	70	34.5	261 2 T11107	cytochrome-c oxida
44	70	34.5	262 2 H71390	cytochrome-c oxida
45	69	34.0	260 2 T11798	cytochrome-c oxida

## ALIGNMENTS

RESULT 1  
OTXL3  
cytochrome-c oxidase (EC 1.9.3.1) chain III - African clawed frog mitochondrion  
N:Alternate names: cytochrome a3 polypeptide III; cytochrome a3 polypeptide III  
C:Species: mitochondrion Xenopus laevis (African clawed frog)  
C:Date: 28-Feb-1986 #sequence\_revision 28-Feb-1986 #text\_change 07-Dec-1999  
C:Accession: A00486  
R:Roel, B.A.; Ma, D.P.; Wilson, R.K.; Wong, J.F.H.  
J. Biol. Chem. 260, 9759-9774, 1985  
A:Title: The complete nucleotide sequence of the Xenopus laevis mitochondrial genome.  
A:Reference number: A00155; M01D:85261388  
A:Accession: A00486  
A:Molecule type: DNA  
A:Residues: 1-260 <ROE>  
A:Cross-references: GB:M10217; GB:X01600; GB:X01601; GB:X02890; NID:g343717; PIDN:AAA  
A:Genetic code: SGC1  
A:Gene: COXIII  
C:Genetics:  
A:Genome: mitochondrion  
A:Superfamily: cytochrome-c oxidase chain III  
C:Keywords: electron transfer; membrane-associated complex; mitochondrial inner membr  
F;16-34/Domain: transmembrane helix #status predicted <TR01>  
F;11-66/Domain: transmembrane helix #status predicted <TR02>  
F;73-105/Domain: transmembrane helix #status predicted <TR04>  
F;129-152/Domain: transmembrane helix #status predicted <TR05>  
F;156-183/Domain: transmembrane helix #status predicted <TR06>  
F;191-223/Domain: transmembrane helix #status predicted <TR07>  
F;233-255/Domain: transmembrane helix #status predicted <TR07>

Query Match 42.9%; Score 87; DB 1; Length 260;  
Best Local Similarity 51.4%; Pred. No. 0.00022;  
Matches 19; Conservative 3; Mismatches 15; Indels 0; Gaps 0;

QY 1 IVHGDRKRTAIGALIAIGLGMIFLCAQAYEYEVHT 37  
DB 151 IMHGDRKRTAIGALIAIGLGMIFLCAQAYEYEVHT 187  
RESULT 2  
T10978  
cytochrome-c oxidase (EC 1.9.3.1) chain III - pig mitochondrion  
C:Species: mitochondrion Sus scrofa domestica (domestic pig)  
C:Date: 16-Jul-1999 #sequence\_revision 16-Jul-1999 #text\_change 20-Jun-2000  
C:Accession: T10978; T11876; C45954  
R:Lin, C.S.; Liu, C.Y.; Sun, Y.L.; Chang, L.C.; Cheng, I.C.; Yang, P.C.; Mo, S.J.T.;  
submitted to the EMBL Data Library, November 1997  
A:Description: Complete nucleotide sequence of the porcine mitochondrial genome.  
A:Reference number: Z17237  
A:Accession: T10978  
A:Status: preliminary; translated from GB/EMBL/DDAJ  
A:Molecule type: DNA

Query Match	41.48;	Score 84;	DB 2;	Length 261;
Best Local Similarity	51.48;	Pred. No. 0.00057;		
Matches 19; Conservative		3; Mismatches 15;		

[illegible]

cytochrome-c oxidase (EC 1.9.3.1) chain III - horse mitochondrion (fragment)  
C.Species: mitochondrion Equus caballus (domestic horse)  
C.Date: 16-Jul-1999 #sequence #revision 16-Jul-1999 #text\_change 13-Aug-1999  
C.Accession: t11863  
R.Xc: Arnason, U.  
Gene 146, 357-362, 1994  
V: The complete mitochondrial DNA sequence of the horse, Equus caballus: Extensive  
Reference number: Z17369, M01D:95047450  
Accession: t11863

A:Status: Preliminary: translated from GB/EMBL/DBD  
A:Molecule type: DNA  
A:Residues: 1-261 <XU>  
A:Cross-references: EMBL:X79547; NID:9577571; PTD:9577578; PIDN:CA56085.1  
C:Genetics:  
C:Genome: mitochondrial  
C:Superfamily: cytochrome-c oxidase chain III  
C:Keywords: electron transfer; memb

Query Match	40.9%	Score 83	DB 2	Length 261
Best Local Similarity	51.4%	Pred. NO	0.00079	
Matches	19	Conservative	3	Mismatches 15
				Indels 0
				Gaps 0
Oy	1	IVHGDRKRTAIGLAIAIGLWIFITLCOAYEYELVHT	37	
		:::>:::	:::>:::	
Db	151	LMEGSRKNNMLOGLEFITTISGLGVFTLLQASEIYEASFT	187	

```
RESULT 4
T11369
cytochrome-c oxidase (EC 1.9.3.1) chain III - donkey mitochondrion (fragment)
C:Species: mitochondrion Equus asinus (donkey)
C:Date: 16-Jul-1999 #sequence_revision 16-Jul-1999 #text_change 21-Jul-2000
C:Accession: T11369
```

R: Xu, X.; Gulberg, A.; Arnason, U.  
J. Mol. Evol. 43: 438-463, 1996  
A: Title: The complete mitochondrial (mtDNA) of the donkey and mtDNA comparisons among  
A: Reference number: Z17265; NID: 97032591  
A: Accession: T11369  
A: Status: preliminary; translated from GB/EMBL/DBJ  
A: Molecule type: DNA  
A: Residues: 1-261 <XNU>  
A: Cross-references: EMBL:X97337; NID: g1805746; PID: CAA66020.1; PID: g1805753  
A: Experimental source: kidney  
C: Genetics: mitochondrial  
A: Genome: mitochondrion  
C: Superfamily: cytochrome-c oxidase chain III  
C: Keywords: electron transfer; membrane-associated complex; mitochondrial inner membr

Query Match 40.9%; Score 83; DB 2; Length 261;  
Best Local Similarity 51.4%; Pred. NO. 0.00079;  
Matches 19; Conservative 1; Mismatches 1

QY	1	IVHDDRKRTAIGIAIIGIGMIFLTLCQAEVVEVYEIT	37	Indels	0;	Gaps	0;
Dd	151	LMEGNRKNNMLOGLEFITSIGVYFFTLQSEVYEAST	187				

RESULT 5  
C45164  
cytochrome-c oxidase (EC 1.9.3.1) chain III - Rhodobacter sphaeroides  
C:Species: Rhodobacter sphaeroides  
C:Date: 26-May-1994 #sequence\_revision 26-May-1994 #text\_change 17-Mar-1999  
C:Accession: C45164  
R:Cao, J.; Hostler, J.; Shapleigh, J.; Revzin, A.; Ferguson-Miller, S.  
J Biol. Chem. 267, 24273-24278, 1992  
A:Title: Cytochrome aa-3 of Rhodobacter sphaeroides

A:Reference number: A45164; MUID:93077508  
A:Accession: C45164  
A:Status: preliminary; not compared with conceptual translation  
A:Molecule type: DNA  
A:Residues: 1-266 <CAO>  
A:Genetics:  
A:Gene: coliI  
A:Superfamily: cytochrome-c oxidase chain III  
A:Keywords: electron transport

```
Query Match      40.6%; Score 82.5; DB 2; Length 266;
Best Local Similarity 56.2%; Pred. No. 0.00094;
Matches 18; Conservative 6; Mismatches 7; Indels 1; Gaps 1;
```

RESULT 6  
T11293  
cytochrome-c oxidase (EC 1.9.3.1) chain III - Raja radiata mitochondrion  
C.Species: mitochondrion Raja radiata  
C.Date: 16-Jul-1999 #sequence\_revision 16-Jul-1999 #text\_change 11-May-2000  
R.Rasmussen, A.S.; Arnason, U.  
Proc. Natl. Acad. Sci. U.S.A. 96, 2177-2182, 1999  
A.Title: Molecular studies suggest that cartilaginous fishes have an apical position  
A.Reference number: 217259; PMID:99162577  
A.Accession: T11293  
A>Status: preliminary; translated from GB/EMBL/DBJ  
A:Molecule type: DNA  
A:Residues: 1-261 <RAS>  
A:Cross-references: EMBL:AF106038; NID:g4406269; PID:g4406276; PIDN:AAD1935.1  
C.Genetics:  
A:Gene: COX-III  
A:Genome: mitochondrion





QY	1	IVHGDRKKTALGALAIAGLWIFTLQAYEYVEIVHT	37
		:   :   :   :   :   :   :   :   :   :   :	
Db	151	IMEGERRKQTIQALTWITLLGFYFTFLQMEYEAFTT	187

Matches	17;	Conservative	5;	Mismatches	15;	Indels	0;	Gaps	0;
1	IYHGDRKKTATGTAIAIGLGMIFLLQAAVEYEIVHT	37							

151 IMAGEKÖTÖLITILLGFFFTFLÖGMEYEAFT 187

cytochrome-c oxidase (EC 1.9.3.1) chain III [validated] - bovine mitochondrion  
 C:Species: mitochondrion Bos primigenius taurus (cattle)  
 C:Date: 18-Aug-1982 #sequence\_revision 18-Aug-1982 #text\_change 15-Sep-2000  
 C:Accession: A00483  
 R:Anderson, S.; de Bruijn, M.H.L.; Coulson, A.R.; Eperon, I.C.; Sanger, F.; Young, I.G.  
 J. Mol. Biol. 156, 683-717, 1982  
 A:Title: Complete sequence of bovine mitochondrial DNA. Conserved features of the mammal  
 A:Reference number: A00152; MUID:83010260  
 A:Accession: A00483  
 A:Molecule type: DNA  
 A:Residues: 1-261 <AND>  
 A:Cross-references: GB:J01394; NID:9336430; EMBL:V00654; NID:912800; PID:912807  
 R:Tsukihara, T.; Aoyama, H.; Yamashita, E.; Tomizaki, T.; Yamaguchi, H.; Shinzawa-Itoh,  
 submitted to the Brookhaven Protein Data Bank, April 1996  
 A:Reference number: A67451; PDB:1OCC  
 A:Contents: annotation: X-ray crystallography, 2.8 angstroms, residues 1-261  
 R:Tsukihara, T.; Aoyama, H.; Yamashita, E.; Tomizaki, T.; Yamaguchi, H.; Shinzawa-Itoh,  
 Science 272, 1136-1144, 1996  
 A:Title: The whole structure of the 13-subunit oxidized cytochrome c oxidase at 2.8 ang  
 A:Reference number: A57981; MUID:96216288  
 A:Contents: annotation: X-ray crystallography, 2.8 angstroms  
 C:Genetics:  
 A:Gene: coxIII  
 A:Genome: mitochondrion  
 A:Genetic code: SGC1  
 C:Complex: part of a 13 chain complex spanning the inner mitochondrial membrane and cons  
 Vb (see PIR:OGBO6A), Via (see PIR:OGBO6), Vtb (see PIR:OGBO7), Vtc (see PIR:OGBO6C), VI  
 imers within the mitochondrial inner-membrane  
 C:Function:  
 A:Description: the cytochrome-c oxidase complex catalyzes the oxidation of four molecule  
 ns from the mitochondrial matrix producing two molecules of water and lowering the conce  
 A:pathway: oxidative phosphorylation: respiratory chain  
 A>Note: chain III may help bind cytochrome c on the intracristal side of the inner-membr  
 C:Superfamily: cytochrome-c oxidase chain III  
 C:Keywords: electron transfer; membrane-associated complex; mitochondrial inner membrane  
 F:1-15/Domain: mitochondrial matrix #status experimental <MM1>  
 F:16-34/Domain: transmembrane helix #status experimental <TR01>  
 F:35-40/Domain: intracristal #status experimental <ITC1>  
 F:41-66/Domain: transmembrane helix #status experimental <TR02>  
 F:67-72/Domain: mitochondrial matrix #status experimental <TR03>  
 F:73-105/Domain: transmembrane helix #status experimental <MM2>  
 F:106-128/Domain: intracristal #status experimental <ITC2>  
 F:129-152/Domain: transmembrane helix #status experimental <TR04>  
 F:153-155/Domain: mitochondrial matrix #status experimental <MM3>  
 F:156-183/Domain: transmembrane helix #status experimental <TR05>  
 F:184-190/Domain: intracristal #status experimental <ITC3>  
 F:191-223/Domain: transmembrane helix #status experimental <TR06>  
 F:224-232/Domain: mitochondrial matrix #status experimental <MM4>  
 F:233-256/Domain: transmembrane helix #status experimental <TR07>  
 F:257-261/Domain: intracristal #status experimental <ITC4>

Query Match 38.9%; Score 79; DB 1; Length 261;  
 Best Local Similarity 51.4%; Pred. No. 0.0028;  
 Matches 19; Conservative 2; Mismatches 16; Indels 0; Gaps 0;

OY 1 IYHSDRKRTAIGLATAIGLWIFTLQAVYEYIYHT 37  
 :: ||||| | | | | | | | | | | | | | | | |  
 Db 151 LMEGDRKMLQALFTITLGVFTLLQASEYYEAPPT 187

Search completed: January 10, 2002, 02:00:24  
 Job time: 547 sec





```

COX3_PIG
ID COX3_PIG STANDARD; PRT; 261 AA.
AC Q35916; 079879;
DT 15-DEC-1998 (Rel. 37, Created)
DT 15-DEC-1998 (Rel. 37, Last sequence update)
DE 30-MAY-2000 (Rel. 39, Last annotation update)
DE CYTOCHROME C OXIDASE POLYPEPTIDE III (EC 1.9.3.1).
GN MTCO3 OR COIII.
OS Sus scrofa (Pig).
OC Mitochondrion.
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Cetartiodactyla; Suidae; Sus.
OX NCBI_TaxID=9823;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=98404150; PubMed=9732457;
RA Drsing B.M., Arnason U.;
RT "The complete mitochondrial DNA sequence of the pig (Sus scrofa).";
RL J. Mol. Evol. 47:302-306(1998).
RN [2]
RP SEQUENCE OF 1-30 FROM N.A.
RX MEDLINE=86295557; PubMed=3017295;
RA Matsubara T., Hayashi Y., Kimura J., Yasuda Y., Saitou N., Tomita T.,
RA Ogasawara N.;
RT "Pig mitochondrial DNA: polymorphism, restriction map orientation,
RT and sequence data.";
RL Biochem. Genet. 24:385-396(1986).
CC -1- FUNCTION: SUBUNIT I, II, AND III FORM THE FUNCTIONAL CORE OF
CC THE ENZYME COMPLEX.
CC -1- CATALYTIC ACTIVITY: 4 FERROCYTOCHROME C + O(2) = 2 H(2)O +
CC 4 FERRICYTOCHROME C.
CC -1- SIMILARITY: BELONGS TO THE CYTOCHROME C OXIDASE SUBUNIT 3 FAMILY.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@sib-sib.ch).
CC -----
DR EMBL: AJ002189; CAA05235.1; ALT_INIT.
DR EMBL: M26139; AAA32032.1; -.
DR InterPro: IPR000298; Cytc_oxdse_III.
DR Pfam: PF00510; COX3; 1.
DR PROSITE: PS50253; COX3; 1.
DR OXidoreductase; Mitochondrion; Transmembrane.
KW CONFLICT 22 24 LSA -> YSG (IN REF. 2).
PT SEQUENCE 261 AA; 29728 MW; 07E93FCA7FC1784B CRC64;
SQ

```

Query Match 41.4%; Score 84; DB 1; Length 261;  
Best Local Similarity 51.4%; Pred. No. 0.00087;  
Matches 19; Conservative 3; Mismatches 15; Indels 0; Gaps 0;

```

OY 1 IVHGDRKRTAIGLAIGMIFTLQAYEYETVHT 37
DB 151 LMEGDRKRMIGALSTITALGYFTLLQASEYEYEAFT 187

```

RESULT 3  
COX3\_EQUUS  
ID COX3\_EQUUS STANDARD; PRT; 261 AA.  
AC P92481;  
DT 15-JUL-1998 (Rel. 36, Created)  
DT 15-JUL-1998 (Rel. 36, Last sequence update)  
DT 15-JUL-1999 (Rel. 38, Last annotation update)  
DE CYTOCHROME C OXIDASE POLYPEPTIDE III (EC 1.9.3.1).  
GN MTCO3 OR COIII.  
OS Equus asinus (Donkey).  
OC Mitochondrion.  
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
OC Mammalia; Eutheria; Perissodactyla; Equidae; Equus.

```

OX NCBI_TaxID=9793;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=97032591; PubMed=8875857;
RA Xu X., Gullberg A., Arnason U.;
RT "The complete mitochondrial DNA (mtDNA) of the donkey and mtDNA
RT comparisons among four closely related mammalian species pairs.";
RL J. Mol. Evol. 43:438-463(1996).
CC -1- FUNCTION: SUBUNIT I, II, AND III FORM THE FUNCTIONAL CORE OF
CC THE ENZYME COMPLEX.
CC -1- CATALYTIC ACTIVITY: 4 FERROCYTOCHROME C + O(2) = 2 H(2)O +
CC 4 FERRICYTOCHROME C.
CC -1- SIMILARITY: BELONGS TO THE CYTOCHROME C OXIDASE SUBUNIT 3 FAMILY.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@sib-sib.ch).
CC -----
DR EMBL: X97337; CAA66020.1; -.
DR InterPro: IPR000298; Cytc_oxdse_III.
DR Pfam: PF00510; COX3; 1.
DR PROSITE: PS50253; COX3; 1.
DR OXidoreductase; Mitochondrion; Transmembrane.
KW SEQUENCE 261 AA; 29761 MW; EB33C3F0BE40769 CRC64;
SQ

```

Query Match 40.9%; Score 83; DB 1; Length 261;  
Best Local Similarity 51.4%; Pred. No. 0.0012;  
Matches 19; Conservative 3; Mismatches 15; Indels 0; Gaps 0;

```

OY 1 IVHGDRKRTAIGLAIGMIFTLQAYEYETVHT 37
DB 151 LMEGDRKRMIGALSTITALGYFTLLQASEYEYEAFT 187

```

RESULT 4  
COX3\_HORSE  
ID COX3\_HORSE STANDARD; PRT; 261 AA.  
AC P48661;  
DT 01-FEB-1996 (Rel. 33, Created)  
DT 01-FEB-1996 (Rel. 33, Last sequence update)  
DT 15-JUL-1999 (Rel. 38, Last annotation update)  
DE CYTOCHROME C OXIDASE POLYPEPTIDE III (EC 1.9.3.1).  
GN MTCO3 OR COIII.  
OS Equus caballus (Horse).  
OC Mitochondrion.  
OC Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;  
OC Mammalia; Eutheria; Perissodactyla; Equidae; Equus.  
OX NCBI\_TaxID=9796;  
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=95047450; PubMed=7958969;
RA Xu X., Arnason U.;
RT "The complete mitochondrial DNA sequence of the horse, Equus
RT caballus: extensive heteroplasmy of the control region.";
RL Gene 148:357-362(1994).
CC -1- FUNCTION: SUBUNIT I, II, AND III FORM THE FUNCTIONAL CORE OF
CC THE ENZYME COMPLEX.
CC -1- CATALYTIC ACTIVITY: 4 FERROCYTOCHROME C + O(2) = 2 H(2)O +
CC 4 FERRICYTOCHROME C.
CC -1- SIMILARITY: BELONGS TO THE CYTOCHROME C OXIDASE SUBUNIT 3 FAMILY.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>



FT TRANSMEM 48 76  
 FT DOMAIN 77 78 CYTOPLASMIC.  
 FT TRANSMEM 79 114  
 FT DOMAIN 115 138 PERIPLASMIC.  
 FT TRANSMEM 139 165  
 FT DOMAIN 166 167 CYTOPLASMIC.  
 FT TRANSMEM 168 196  
 FT DOMAIN 197 202 PERIPLASMIC.  
 FT TRANSMEM 203 236  
 FT DOMAIN 237 243 CYTOPLASMIC.  
 FT TRANSMEM 244 273  
 SEQUENCE 273 AA; 30655 MW; 0831BD966AE3C7D6 CRC64;

Query Match 39.4%; Score 81; DB 1; Length 273;  
 Best Local Similarity 54.8%; Pred. No. 0.0022;  
 Matches 17; Conservative 4; Mismatches 10; Indels 0; Gaps 0;  
 OY 1 IVHGRKRTAIGLAIGMIFTLQAYEYEVHT 31  
 Db 163 VLEGRKRTTNGLVAVILGVCFTGLQAYEV 193

RESULT 7  
 ID COX3\_ONCCL STANDARD; PRT; 261 AA.  
 AC P20684;  
 DT 01-FEB-1991 (Rel. 17, Created)  
 DT 01-FEB-1991 (Rel. 17, Last sequence update)  
 DE 15-JUL-1999 (Rel. 38, Last annotation update)  
 DE CYTOCHROME C OXIDASE POLYPEPTIDE III (EC 1.9.3.1).  
 GN COXIII OR COIII.  
 OS Oncorhynchus clarki (Cutthroat trout) (Salmo clarki).  
 OC Mitochondrion.  
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 OC Actinopterygii; Neopterygii; Teleostei; Euteleostei;  
 OC Protacanthopterygii; Salmoniformes; Salmonidae; Oncorhynchus.  
 OX NCBI\_TaxID=30962;  
 RN [1]  
 KW OXidoreductase; Mitochondrion; Transmembrane.  
 SEQUENCE FROM N.A.  
 RX MEDLINE=89382658; PubMed=2550657;  
 RA Thomas W.K.; Beckenbach A.T.;  
 RT "Variation in salmonid mitochondrial DNA: evolutionary constraints  
 RL J. Mol. Evol. 29:233-245(1989)."  
 CC -1- FUNCTION: SUBUNIT I, II, AND III FORM THE FUNCTIONAL CORE OF  
 CC THE ENZYME COMPLEX.  
 CC -1- CATALYTIC ACTIVITY: 4 FERROCYTOCHROME C + O(2) = 2 H(2)O +  
 CC 4 FERRICCYTOCHROME C.  
 CC -1- SIMILARITY: BELONGS TO THE CYTOCHROME C OXIDASE SUBUNIT 3 FAMILY.  
 DR HSSP: P00415; IOCC.  
 DR InterPro: IPR000298; Cytc\_oxdse\_III.  
 DR Pfam: PF00510; COX3; 1.  
 DR PROSITE: PS0253; COX3; 1.  
 KW OXidoreductase; Mitochondrion; Transmembrane.  
 SEQUENCE 261 AA; 29705 MW; 434ECAA20954625E CRC64;

Query Match 39.4%; Score 80; DB 1; Length 261;  
 Best Local Similarity 45.9%; Pred. No. 0.0028;  
 Matches 17; Conservative 5; Mismatches 15; Indels 0; Gaps 0;  
 OY 1 IVHGRKRTAIGLAIGMIFTLQAYEYEVHT 37  
 Db 151 IMEGRKRTTNGLVAVILGVCFTGLQAYEYEVHT 187

RESULT 8  
 COX3\_ONCCL STANDARD; PRT; 261 AA.  
 ID COX3\_ONCCL  
 AC P48172;  
 DT 01-FEB-1996 (Rel. 33, Created)

DT 01-FEB-1996 (Rel. 33, Last sequence update)  
 DT 15-JUL-1999 (Rel. 38, Last annotation update)  
 DE CYTOCHROME C OXIDASE POLYPEPTIDE III (EC 1.9.3.1).  
 GN COXIII OR COIII.  
 OS Oncorhynchus mykiss (Rainbow trout) (Salmo gairdneri).  
 OC Mitochondrion.  
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 OC Actinopterygii; Neopterygii; Teleostei; Euteleostei;  
 OC Protacanthopterygii; Salmoniformes; Salmonidae; Oncorhynchus.  
 OX NCBI\_TaxID=8022;  
 RN [1]  
 KW OXidoreductase; Mitochondrion; Transmembrane.  
 SEQUENCE FROM N.A.  
 RP TISSUE=Liver;  
 RX MEDLINE=96139027; PubMed=8587139;  
 RA Zardoya R.; Garrido-Perterra A.; Baulista J.M.;  
 RT "The complete nucleotide sequence of the mitochondrial DNA genome of  
 RT the rainbow trout, *Oncorhynchus mykiss*.",  
 RL J. Mol. Evol. 41:942-951(1995).  
 CC -1- FUNCTION: SUBUNIT I, II, AND III FORM THE FUNCTIONAL CORE OF  
 CC THE ENZYME COMPLEX.  
 CC -1- CATALYTIC ACTIVITY: 4 FERROCYTOCHROME C + O(2) = 2 H(2)O +  
 CC 4 FERRICCYTOCHROME C.  
 CC -1- SIMILARITY: BELONGS TO THE CYTOCHROME C OXIDASE SUBUNIT 3 FAMILY.  
 CC THIS SWISS-PROT entry is copyright. It is produced through a collaboration  
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -  
 CC the European Bioinformatics Institute. There are no restrictions on its  
 CC use by non-profit institutions as long as its content is in no way  
 CC modified and this statement is not removed. Usage by and for commercial  
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
 CC or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).  
 DR EMBL: L29771; AAB03353.1;  
 DR HSSP: P00415; IOCC.  
 DR InterPro: IPR000298; Cytc\_oxdse\_III.  
 DR Pfam: PF00510; COX3; 1.  
 DR PROSITE: PS0253; COX3; 1.  
 KW OXidoreductase; Mitochondrion; Transmembrane.  
 SEQUENCE 261 AA; 29675 MW; ADAF7AA21255D96E CRC64;

Query Match 39.4%; Score 80; DB 1; Length 261;  
 Best Local Similarity 45.9%; Pred. No. 0.0028;  
 Matches 17; Conservative 5; Mismatches 15; Indels 0; Gaps 0;  
 OY 1 IVHGRKRTAIGLAIGMIFTLQAYEYEVHT 37  
 Db 151 IMEGRKRTTNGLVAVILGVCFTGLQAYEYEVHT 187

RESULT 9  
 COX3\_ONCCL STANDARD; PRT; 261 AA.  
 ID COX3\_ONCCL  
 AC P20683;  
 DT 01-FEB-1991 (Rel. 17, Created)  
 DT 01-FEB-1991 (Rel. 17, Last sequence update)  
 DE 15-JUL-1999 (Rel. 38, Last annotation update)  
 DE CYTOCHROME C OXIDASE POLYPEPTIDE III (EC 1.9.3.1).  
 GN COXIII OR COIII.  
 OS Oncorhynchus mykiss (Rainbow trout) (Salmo gairdneri).  
 OC Mitochondrion.  
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 OC Actinopterygii; Neopterygii; Teleostei; Euteleostei;  
 OC Protacanthopterygii; Salmoniformes; Salmonidae; Oncorhynchus.  
 OX NCBI\_TaxID=8023; 8020;  
 RN [1]  
 KW OXidoreductase; Mitochondrion; Transmembrane.  
 SEQUENCE FROM N.A.  
 RP SPECIES=O. mykiss;  
 RX MEDLINE=89382658; PubMed=2550657;  
 RA Thomas W.K.; Beckenbach A.T.;  
 RT "Variation in salmonid mitochondrial DNA: evolutionary constraints  
 RT and mechanisms of substitution.",



```

RL J. Mol. Evol. 29:233-245(1989).
RN
RP SEQUENCE FROM N.A.
RC SPECIES=O. masou;
RC MEDLINE=96261239; PubMed=8688811;
RA Oohara I., Okazaki T.;
RT "Genetic relationship among three subspecies of Oncorhynchus masou
determined by mitochondrial DNA sequence analysis.";
RL Zool. Sci. 13:189-198(1996).
CC -1- FUNCTION: SUBUNIT I, II, AND III FORM THE FUNCTIONAL CORE OF
CC THE ENZYME COMPLEX.
CC -1- CATALYTIC ACTIVITY: 4 FERROCYTOCHROME C + O(2) = 2 H(2)O +
CC 4 FERROCYTOCHROME C.
CC -1- SIMILARITY: BELONGS TO THE CYTOCHROME C OXIDASE SUBUNIT 3 FAMILY.
CC
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch)
CC -----
CC EMBL: D63410; BAA09711.1; -
CC EMBL: D63336; BAA09653.1; -
CC PIR: A30396; A30396.
CC PIR: B30396; B30396.
CC PIR: H34012; H34012.
CC HSSP: P00415; 10CC.
CC InterPro: IPR000298; Cytc_oxdse_III.
CC Pfam: PF00510; COX3; 1.
CC PROSITE: PS50253; COX3; 1.
CC Oxidoreductase: Mitochondrion; Transmembrane.
KM SEQUENCE 261 AA; 29691 MW; 362F361768AD0635 CRC64;
SQ
Query Match 39.4%; Score 80; DB 1; Length 261;
Best Local Similarity 45.9%; Pred. No. 0.0028;
Matches 17; Conservative 5; Mismatches 15; Indels 0; Gaps 0;
OY 1 IVHGDKRTATGAIATGAGTFLQAYEYEVHT 37
|:|:|:| | | | | | | | | | | | | | |
Db 151 IMEGRKOTIQLVTLTLLGFYFPLQMEYEAFT 187
RESULT 10
COX3-SALSA STANDARD: PRT; 261 AA.
AC Q36860; 003380;
DT 01-NOV-1997 (Rel. 35, Created)
DT 30-MAY-2000 (Rel. 39, Last sequence update)
DT 20-AUG-2001 (Rel. 40, Last annotation update)
DE CYTOCHROME C OXIDASE POLYPEPTIDE III (EC 1.9.3.1).
GN COXIII OR COXIII.
OS Salmo salar (Atlantic salmon).
OG Mitochondrion.
OC Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;
OC Actinopterygii; Neopterygii; Teleostei; Euteleostei;
OC Protacanthopterygii; Salmoniformes; Salmonidae; Salmo.
OX NCBI_TaxID=8030;
RN
RP SEQUENCE FROM N.A.
RC TISSUE=Liver;
RC MEDLINE=95093530; PubMed=8000479;
RA Hardman G., Byrnes L., Peden J., Wolff J., Gannon F.;
RT "Cloning and sequencing of the Atlantic salmon (Salmo salar)
cytochrome c oxidase subunit III gene (coxiii) and analysis of coxiii
expression during parv-smolt transformation.";
RT Mol. Mar. Biol. Biotechnol. 3:210-216(1994).
RN
RP SEQUENCE FROM N.A.
RC TISSUE=Kidney;

```

```

RA Hardman G.T., Wolff J., Peden J., Gannon F.;
RT "Isolation of Atlantic Salmon (Salmo salar) cytochrome c oxidase
subunit II gene (coxii).";
RL J. Appl. Ichthyol. 10:64-68(1994).
RN
RP SEQUENCE FROM N.A.
RC MEDLINE=97159563; PubMed=9007022;
RA Oohara I., Sawano K., Okazaki T.;
RT "Mitochondrial DNA sequence analysis of the masu salmon -- phylogeny
in the genus Oncorhynchus.";
RL Mol. Phylogenet. Evol. 7:71-78(1997).
RN
RP SEQUENCE FROM N.A.
RC TISSUE=Liver;
RC MEDLINE=20018174; PubMed=10548724;
RA Hurst C.D., Bartlett S.E., Davidson W.S., Bruce I.J.;
RT "The complete mitochondrial DNA sequence of the Atlantic salmon, Salmo
salar.";
RL Gene 239:237-242(1999).
RN
RP SEQUENCE FROM N.A.
RA Arnason U., Jonsson E., Rasmussen A.S.;
RT "The complete mitochondrial genome sequence of a teleost, Salmo salar,
and comparisons with other salmoniformes.";
RL Submitted (MAR-1999) to the EMBL/Genbank/DBJ databases.
CC -1- FUNCTION: SUBUNIT I, II, AND III FORM THE FUNCTIONAL CORE OF
CC THE ENZYME COMPLEX.
CC -1- CATALYTIC ACTIVITY: 4 FERROCYTOCHROME C + O(2) = 2 H(2)O +
CC 4 FERROCYTOCHROME C.
CC -1- SIMILARITY: BELONGS TO THE CYTOCHROME C OXIDASE SUBUNIT 3 FAMILY.
CC
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch)
CC -----
CC EMBL: L04502; AAA62409.1; -
CC EMBL: D84148; BAA20158.1; -
CC EMBL: U12143; AAD04739.1; -
CC EMBL: AF133701; AAF61384.1; -
CC InterPro: IPR000298; Cytc_oxdse_III.
CC Pfam: PF00510; COX3; 1.
CC PROSITE: PS50253; COX3; 1.
CC Oxidoreductase: Mitochondrion; Transmembrane.
KM CONFLICT 135 A -> S (IN REF. 1 AND 2).
FT CONFLICT 158 Q -> R (IN REF. 1 AND 2).
FT CONFLICT 237 A -> L (IN REF. 1 AND 2).
SQ SEQUENCE 261 AA; 29795 MW; 0E11AF391B93C55A CRC64;
Query Match 39.4%; Score 80; DB 1; Length 261;
Best Local Similarity 45.9%; Pred. No. 0.0028;
Matches 17; Conservative 5; Mismatches 15; Indels 0; Gaps 0;
OY 1 IVHGDKRTATGAIATGAGTFLQAYEYEVHT 37
|:|:|:| | | | | | | | | | | | | | |
Db 151 IMEGRKOTIQLVTLTLLGFYFPLQMEYEAFT 187
RESULT 11
COX3-BOVIN STANDARD: PRT; 261 AA.
AC P00415;
DT 21-JUL-1986 (Rel. 01, Created)
DT 21-JUL-1986 (Rel. 01, Last sequence update)
DT 20-AUG-2001 (Rel. 40, Last annotation update)
DE CYTOCHROME C OXIDASE POLYPEPTIDE III (EC 1.9.3.1).
GN MTXCO3 OR COXIII.
OS Bos taurus (Bovine).
OG Mitochondrion.

```

OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 OC Mammalia; Eutheria; Cetartiodactyla; Ruminantia; Pecora; Bovidae;  
 OC Bovidae; Bos.  
 OC NCBI\_TaxID=9913;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RC TISSUE-Heart;  
 RA MEDLINE=83010260; PubMed=7120390;  
 RA Anderson S., de Bruijn M.H.L., Coulson A.R., Eperon I.C., Sanger F.,  
 RA Young I.G.;  
 RT "Complete sequence of bovine mitochondrial DNA. Conserved features of  
 RT the mammalian mitochondrial genome.";  
 RL J. Mol. Biol. 156:683-717(1982).  
 RN [2]  
 RP X-RAY CRYSTALLOGRAPHY (2.8 ANGSTROMS).  
 RX MEDLINE=96216288; PubMed=8638158;  
 RA Tsukihara T., Aoyama H., Yamashita E., Tomizaki T., Yamaguchi H.,  
 RA Shinzawa-Itoh K., Nakashima R., Yaono R., Yoshikawa S.;  
 RT "The whole structure of the 13-subunit oxidized cytochrome c oxidase  
 RT at 2.8 A.";  
 RL Science 272:1136-1144(1996).  
 RN [3]  
 RP X-RAY CRYSTALLOGRAPHY (2.8 ANGSTROMS).  
 RC TISSUE-Heart;  
 RX MEDLINE=99190827; PubMed=10089392.  
 RA Tomizaki T., Yamashita E., Yamaguchi H., Aoyama H., Tsukihara T.,  
 RA Shinzawa-Itoh K., Nakashima R., Yaono R., Yoshikawa S.;  
 RT "Structure analysis of bovine heart cytochrome c oxidase at 2.8 A  
 RT resolution.";  
 RL Acta Crystallogr. D 55:31-45(1999).  
 RN [4]  
 RP X-RAY CRYSTALLOGRAPHY (2.9 ANGSTROMS).  
 RC TISSUE-Heart;  
 RX MEDLINE=20235645; PubMed=10771420;  
 RA Fei M.J., Yamashita E., Inoue N., Yao M., Yamaguchi H., Tsukihara T.,  
 RA Shinzawa-Itoh K., Nakashima R., Yoshikawa S.;  
 RT "X-ray structure of azide-bound fully oxidized cytochrome c oxidase  
 RT from bovine heart at 2.9 A resolution.";  
 RL Acta Crystallogr. D 56:529-535(2000).  
 CC -1- FUNCTION: SUBUNIT I, II, AND III FORM THE FUNCTIONAL CORE OF  
 CC THE ENZYME COMPLEX.  
 CC -1- CATALYTIC ACTIVITY: 4 FERROCYTOCHROME C + O(2) = 2 H(2)O +  
 CC 4 FERROCYTOCHROME C.  
 CC -1- SUBCELLULAR LOCATION: INTEGRAL MEMBRANE PROTEIN. MITOCHONDRIAL  
 CC INNER MEMBRANE.  
 CC -1- SIMILARITY: BELONGS TO THE CYTOCHROME C OXIDASE SUBUNIT 3 FAMILY.  
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -  
 CC the European Bioinformatics Institute. There are no restrictions on its  
 CC use by non-profit institutions as long as its content is in no way  
 CC modified and this statement is not removed. Usage by and for commercial  
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
 CC or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).  
 CC -----  
 DR EMBL: V00654; CAA24003.1; ALT\_TERM.  
 DR EMBL: J01394; AAB59274.1; -.  
 DR PIR: A00483; OTB03.  
 DR PDB: 1OCC; 07-DEC-96.  
 DR PDB: 2OCC; 26-MAY-98.  
 DR PDB: 1OCR; 29-JUL-99.  
 DR PDB: 1OCO; 22-JUL-99.  
 DR PDB: 1OCC; 22-JUL-99.  
 DR InterPro: IPR000298; Cytc\_oxdse\_III.  
 DR Pfam: PF00510; COX3; 1.  
 DR PROSITE: PS50253; COX3; 1.  
 DR Oidoreductase; Mitochondrion; Transmembrane; 3D-structure;  
 KM Inner membrane.  
 FT TRANSMEM 1 15 MITOCHONDRIAL MATRIX.  
 FT DOMAIN 16 34 I.  
 FT DOMAIN 35 40 MITOCHONDRIAL INTERMEMBRANE.  
 FT TRANSMEM 41 66 II.  
 FT DOMAIN 67 72 MITOCHONDRIAL MATRIX.

FT TRANSMEM 73 105 III.  
 FT DOMAIN 106 128 MITOCHONDRIAL INTERMEMBRANE.  
 FT TRANSMEM 129 152 IV.  
 FT DOMAIN 153 155 MITOCHONDRIAL MATRIX.  
 FT TRANSMEM 156 183 V.  
 FT DOMAIN 184 190 MITOCHONDRIAL INTERMEMBRANE.  
 FT TRANSMEM 191 223 VI.  
 FT DOMAIN 224 232 MITOCHONDRIAL MATRIX.  
 FT TRANSMEM 233 256 VII.  
 FT DOMAIN 257 261 MITOCHONDRIAL INTERMEMBRANE.  
 SO SEQUENCE 261 AA; 29919 MW; 0547E76A9089B3F8 CRC64;  
 Query Match 38.9%; Score 79; DB 1; Length 261;  
 Best Local Similarity 51.4%; Pred. No. 0.0038;  
 Matches 19; Conservative 2; Mismatches 16; Indels 0; Gaps 0;  
 Oy 1 IVHGDKRRTAIGAIAGIIGWITLCQAYEYVEIYHT 37  
 Db 151 LMEDRRKMLQALFTITLTGVTTLQASEYVEAPET 187  
 RESULT 12  
 COX3\_CYPCA  
 ID COX3\_CYPCA STANDARD: PRT; 261 AA.  
 AC P15952;  
 DT 01-APR-1990 (Rel. 14, Created)  
 DT 01-FEB-1994 (Rel. 28, Last sequence update)  
 DT 15-JUL-1999 (Rel. 38, Last annotation update)  
 DE CYTOCHROME C OXIDASE POLYPEPTIDE III (EC 1.9.3.1).  
 GN COXIII OR COXII.  
 OS Cyprinus carpio (Common carp).  
 OC Mitochondrion.  
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 OC Actinopterygii; Neopterygii; Teleostei; Euteleostei; Ostariophysi;  
 OC Cypriniformes; Cyprinidae; Cyprininae; Cyprinus.  
 OC NCBI\_TaxID=7962;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RX MEDLINE=90192097; PubMed=2156223;  
 RA Huang C.J., Huang F.L., Chang Y.S., Tsai Y.J., Lo T.B.;  
 RT "Nucleotide sequence of carp mitochondrial cytochrome c oxidase III.";  
 RL Nucleic Acids Res. 18:1056-1056(1990).  
 RN [2]  
 RP SEQUENCE FROM N.A.  
 RX MEDLINE=94223691; PubMed=8169959;  
 RA Chang Y.S., Huang F.L., Lo T.B.;  
 RT "The complete nucleotide sequence and gene organization of carp  
 RT (Cyprinus carpio) mitochondrial genome.";  
 RL J. Mol. Evol. 38:138-155(1994).  
 CC -1- FUNCTION: SUBUNIT I, II, AND III FORM THE FUNCTIONAL CORE OF  
 CC THE ENZYME COMPLEX.  
 CC -1- CATALYTIC ACTIVITY: 4 FERROCYTOCHROME C + O(2) = 2 H(2)O +  
 CC 4 FERROCYTOCHROME C.  
 CC -1- SIMILARITY: BELONGS TO THE CYTOCHROME C OXIDASE SUBUNIT 3 FAMILY.  
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -  
 CC the European Bioinformatics Institute. There are no restrictions on its  
 CC use by non-profit institutions as long as its content is in no way  
 CC modified and this statement is not removed. Usage by and for commercial  
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
 CC or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).  
 CC -----  
 DR EMBL: X17006; CAA34866.1; -.  
 DR EMBL: X61010; CAA43341.1; -.  
 DR PIR: S36010; OTC43.  
 DR HSSP: P00415; 1OCC.  
 DR InterPro: IPR000298; Cytc\_oxdse\_III.  
 DR Pfam: PF00510; COX3; 1.  
 DR PROSITE: PS50253; COX3; 1.  
 DR Oidoreductase; Mitochondrion; Transmembrane.  
 FT CONFLICT 121 121 I -> M (IN REF. 1).

SQ SEQUENCE 261 AA: 29640 MW: 285D890751412EP5 CRC64;

Query Match 38.9%; Score 79; DB 1; Length 261;  
Best Local Similarity 48.6%; Pred. NO. 0.0038;  
Matches 18; Conservative 4; Mismatches 15; Indels 0; Gaps 0;

QY 1 IYHGDKRRTAIGLAIATIGLWFTLCQAYEYEVHT 37  
151 IMEGNRKQAIQSLATITLGLYFTLQAMEYEAFT 187

RESULT 13

ID COX3\_RABIT STANDARD: PRT: 261 AA.

AC 079433;  
DT 15-JUL-1999 (Rel. 38, Created)  
DT 15-JUL-1999 (Rel. 38, Last sequence update)  
DE CYTOCHROME C OXIDASE POLYPEPTIDE III (EC 1.9.3.1).  
GN MTCO3 OR COIII.  
OS Oryctolagus cuniculus (Rabbit).  
OG Mitochondrion.  
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
OC Mammalia; Eutheria; Lagomorpha; Leporidae; Oryctolagus.  
OX NCBI\_TaxID=9986;  
RN [1]  
RP SEQUENCE FROM N.A.  
RX MEDLINE=98317530; PubMed=9653643;  
RA Gissi C., Gullberg A., Arnason U.;  
RT "The complete mitochondrial DNA sequence of the rabbit, Oryctolagus cuniculus".  
RL Genomics 50:161-169(1998).  
CC -1- FUNCTION: SUBUNIT I, II, AND III FORM THE FUNCTIONAL CORE OF THE ENZYME COMPLEX.  
CC -1- CATALYTIC ACTIVITY: 4 FERROCYTOCHROME C + O(2) = 2 H(2)O +  
CC -1- SIMILARITY: BELONGS TO THE CYTOCHROME C OXIDASE SUBUNIT 3 FAMILY.  
CC This SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL outstation - the European Bioinformatics Institute. There are no restrictions on its use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (See <http://www.isb-sib.ch/announce/> or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).  
CC  
CC EMBL: AJ001588; CAA04853.1;  
DR InterPro: IPR000298; CytC\_oxdse\_III.  
DR Pfam: PF00510; COX3; 1.  
DR PROSITE: PS50253; COX3; 1.  
KM Oxidoreductase; Mitochondrion; Transmembrane.  
SQ SEQUENCE 261 AA: 29730 MW: F32E3CE026201C3F CRC64;

Query Match 38.9%; Score 79; DB 1; Length 261;  
Best Local Similarity 51.4%; Pred. NO. 0.0038;  
Matches 19; Conservative 3; Mismatches 15; Indels 0; Gaps 0;

QY 1 IYHGDKRRTAIGLAIATIGLWFTLCQAYEYEVHT 37  
151 LMENRKNMQALATITLGLYFTLQASEYETST 187

RESULT 14

ID COX3\_DIDMA STANDARD: PRT: 261 AA.

AC P41312;  
DT 01-FEB-1995 (Rel. 31, Created)  
DT 01-FEB-1995 (Rel. 31, Last sequence update)  
DT 15-JUL-1999 (Rel. 38, Last annotation update)  
DE CYTOCHROME C OXIDASE POLYPEPTIDE III (EC 1.9.3.1).  
GN MTCO3 OR COIII.

OS Didelphis marsupialis virginiana (North American opossum).

OG Mitochondrion.  
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
OC Mammalia; Metatheria; Didelphimorphia; Didelphidae; Didelphis.

OX NCBI\_TaxID=9267;

RN [1]

RP SEQUENCE FROM N.A.

RX TISSUE=Liver;  
MEDLINE=9433786; PubMed=8056314;

RA Janke A., Feldmaier-Fuchs G., Thomas K., von Haeseler A., Pabo S.;  
RT "The marsupial mitochondrial genome and the evolution of placental mammals".

RL Genetids 137:243-256(1994).

CC -1- FUNCTION: SUBUNIT I, II, AND III FORM THE FUNCTIONAL CORE OF THE ENZYME COMPLEX.

CC -1- CATALYTIC ACTIVITY: 4 FERROCYTOCHROME C + O(2) = 2 H(2)O +  
CC -1- SIMILARITY: BELONGS TO THE CYTOCHROME C OXIDASE SUBUNIT 3 FAMILY.

CC This SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL outstation - the European Bioinformatics Institute. There are no restrictions on its use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (See <http://www.isb-sib.ch/announce/> or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).  
CC  
CC EMBL: Z29573; CAA82683.1;  
DR PIR: S42816; S42816.  
DR HSSP: P00415; IOCC.  
DR InterPro: IPR000298; CytC\_oxdse\_III.  
DR Pfam: PF00510; COX3; 1.  
DR PROSITE: PS50253; COX3; 1.  
KM Oxidoreductase; Mitochondrion; Transmembrane.  
SQ SEQUENCE 261 AA: 32135 MW: 51BF64662ECDEAE7 CRC64;

Query Match 38.9%; Score 79; DB 1; Length 261;  
Best Local Similarity 45.9%; Pred. NO. 0.0041;  
Matches 17; Conservative 5; Mismatches 15; Indels 0; Gaps 0;

QY 1 IYHGDKRRTAIGLAIATIGLWFTLCQAYEYEVHT 37  
151 LMENRKNMQALATITLGLYFTLQAMEYEAFT 187

RESULT 15

ID COX3\_BALMU STANDARD: PRT: 261 AA.

AC P41295;  
DT 01-FEB-1995 (Rel. 31, Created)  
DT 01-FEB-1995 (Rel. 31, Last sequence update)  
DT 15-JUL-1999 (Rel. 38, Last annotation update)  
DE CYTOCHROME C OXIDASE POLYPEPTIDE III (EC 1.9.3.1).  
GN MTCO3 OR COIII.  
OS Balanoptera musculus (Blue whale).  
OG Mitochondrion.  
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
OC Mammalia; Eutheria; Cetartiodactyla; Cetacea; Mysticeti;  
OC Balanopteridae; Balanoptera.  
OX NCBI\_TaxID=9771;

RN [1]  
RP SEQUENCE FROM N.A.  
RX MEDLINE=94141932; PubMed=8308901;

RA Arnason U., Gullberg A.;  
RT "Comparison between the complete mtDNA sequences of the blue and the fin whale, two species that can hybridize in nature".

RL J. Mol. Evol. 37:312-322(1993).

CC -1- FUNCTION: SUBUNIT I, II, AND III FORM THE FUNCTIONAL CORE OF THE ENZYME COMPLEX.

CC -1- CATALYTIC ACTIVITY: 4 FERROCYTOCHROME C + O(2) = 2 H(2)O +  
CC -1- SIMILARITY: BELONGS TO THE CYTOCHROME C OXIDASE SUBUNIT 3 FAMILY.



GenCore version 4.5  
Copyright (c) 1993 - 2000 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: January 10, 2002, 02:00:28 : Search time 32.55 Seconds

(Without alignments)  
170.763 Million cell updates/sec

Title: US-09-712-768-6

Perfect score: 203

Sequence: 1 IVHGDRKKTALGALAIAGLWIFTLCOAYEYEVHT 38

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 473505 seqs, 146272329 residues

Total number of hits satisfying chosen parameters: 473505

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :  
1: SP:REMBL\_17:\*  
2: SP:archaea:\*  
3: SP:bacteria:\*  
4: SP:fungi:\*  
5: SP:human:\*  
6: SP\_invertebrate:\*  
7: SP\_mammal:\*  
8: SP\_mhc:\*  
9: SP\_organelle:\*  
10: SP\_plant:\*  
11: SP\_rodent:\*  
12: SP\_virus:\*  
13: SP\_vertebrate:\*  
14: SP\_unclassified:\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

## SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	84	41.4	261	8	Q9MET0 physeter ca
2	84	41.4	267	8	Q9G7T0 sus scrofa
3	84	41.4	267	8	Q9G7S0 sus scrofa
4	84	41.4	278	8	Q9G290 sus scrofa
5	84	41.4	305	8	Q9T244 physcopthor
6	83	40.9	261	8	Q9B629 anguilla ja
7	82	40.4	260	8	Q9GAU2 rana sylvat
8	82	40.4	261	8	Q9B614 carassius a
9	81	39.9	261	8	Q9ZY39 raja radiat
10	81	39.9	261	8	Q9XN34 salvelinus
11	81	39.9	261	8	Q9XN26 salvelinus
12	81	39.9	261	8	Q9G607 diplophos t
13	81	39.9	261	8	Q9G387 echinops te
14	81	39.9	261	8	Q9B263 isodon mac
15	80	39.4	261	8	Q03378 oncorhynch
16	80	39.4	261	8	Q35269 oncorhynch
17	80	39.4	261	8	Q9B603 pleocoglossu
18	79	38.9	261	8	Q9G6F4 polymixia j
19	79	38.9	290	2	Q9A304 caulobacter

20	78	38.4	261	8	Q9TEG5	Q9TEG5 cavia porce
21	77	37.9	109	8	Q99872	Q99872 crotophytus
22	77	37.9	109	8	Q99873	Q99873 crotophytus
23	77	37.9	261	8	Q9MTV4	Q9MTV4 brachydanio
24	77	37.9	261	8	Q9MEV4	Q9MEV4 sciurus vul
25	76	37.4	262	8	Q9TBV9	Q9TBV9 acropora te
26	75	36.9	261	8	Q9XLO8	Q9XLO8 falco pereg
27	75	36.9	261	8	Q9MR64	Q9MR64 coregonus l
28	75	36.9	261	8	Q9G6R8	Q9G6R8 proteopterus
29	74	36.5	261	8	Q95418	Q95418 artibeus ja
30	74	36.5	261	8	Q99600	Q99600 turnix vari
31	74	36.5	261	8	Q9TBG8	Q9TBG8 eudromia el
32	74	36.5	261	8	Q9B6S8	Q9B6S8 ioxodonta a
33	74	36.5	262	8	Q9TA23	Q9TA23 gambella w1
34	73	36.0	128	8	Q9T593	Q9T593 gambella w1
35	73	36.0	261	8	Q9TBG7	Q9TBG7 chaetura pe
36	73	36.0	261	8	Q9TBG5	Q9TBG5 chordelles
37	73	36.0	261	8	Q9MJB0	Q9MJB0 talpa europ
38	73	36.0	261	8	Q9G6T1	Q9G6T1 sardnops m
39	73	36.0	261	8	Q9B909	Q9B909 aulopus jap
40	72	35.5	76	8	Q48210	Q48210 drosophila
41	72	35.5	99	8	Q99876	Q99876 gambella w1
42	72	35.5	99	8	Q9T592	Q9T592 gambella w1
43	72	35.5	128	8	Q9T596	Q9T596 gambella w1
44	72	35.5	138	8	Q9XN56	Q9XN56 drosophila
45	72	35.5	138	8	Q9XK08	Q9XK08 drosophila

## ALIGNMENTS

RESULT 1  
ID Q9MET0 PRELIMINARY: PRT: 261 AA.  
AC Q9MET0:  
DT 01-OCT-2000 (TREMBLrel. 15, Created)  
DT 01-OCT-2000 (TREMBLrel. 15, Last sequence update)  
DT 01-JUN-2001 (TREMBLrel. 17, Last annotation update)  
DE CYTOCHROME OXIDASE SUBUNIT III.  
GN COIT1.  
OS Physeter catodon (Sperm whale) (Physeter macrocephalus).  
OG Mitochondrion.  
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
OC Mammalia; Eutheria; Cetartiodactyla; Cetacea; Odontoceti;  
OC Physeteridae; Physeter.  
OX NCBI\_TaxID=9755;  
RN [1]  
RP SEQUENCE FROM N.A.  
RX MEDLINE=20296818; PubMed=10835487;  
RA Arnason U., Gulberg A., Gretaerdotir S., Ursing B., Janke A.;  
RT "The mitochondrial genome of the sperm whale and a new molecular  
RT reference for estimating eutherian divergence dates.";  
RL J. Mol. Evol. 50:569-578(2000).  
DR EMBL: A0277029; CAB96276.1; .  
DR InterPro: IPR000298; Cytochrome\_oxidase\_III.  
DR Pfam: PF00510; COX3; 1.  
DR PROSITE: PS50253; COX3; 1.  
KW Mitochondrion.  
SQ SEQUENCE 261 AA: 29907 MW: B7010A5A0AF2F28B CRC64;

Query Match 41.4%; Score 84; DB 8; Length 261;  
Best Local Similarity 51.4%; Pred. No. 0.0023;  
Matches 19; Conservative 3; Mismatches 15; Indels 0; Gaps 0;

QY 1 IVHGDRKKTALGALAIAGLWIFTLCOAYEYEVHT 37  
DB 151 LWGDRKQMLALFTTALGLVFTLLQASEYVPTPT 187  
RESULT 2  
Q9G7T0 PRELIMINARY: PRT: 267 AA.

AC Q9G7T0: 01-MAR-2001 (TREMBlrel. 16, Created)  
 DT 01-MAR-2001 (TREMBlrel. 16, Last sequence update)  
 DE 01-JUN-2001 (TREMBlrel. 17, Last annotation update)  
 GN COI1.  
 OS Sus scrofa (Pig).  
 OG Mitochondrion.  
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 OC Mammalia; Eutheria; Cetartiodactyla; Suina; Suidae; Sus.  
 OX NCBI\_TaxID=9823;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RA Kijas J.M.H., Andersson L.;  
 RT "A phylogenetic study of the origin of the domestic pig estimated from  
 the near complete mtDNA genome."  
 RL Submitted (SEP-2000) to the EMBL/GenBank/DBJ databases.  
 DR EMBL; AF304201; AAC28196.1;  
 DR InterPro; IPR000298; CytC\_oxdse\_III.  
 DR Pfam; PF00510; COX3; 1.  
 DR PROSITE; PS50253; COX3; 1.  
 KW Mitochondrion.  
 SQ SEQUENCE 267 AA; 30462 MW; 95C558C54F6C0E3F CRC64;

Query Match 41.4%; Score 84; DB 8; Length 267;  
 Best Local Similarity 51.4%; Pred. No. 0.0023;  
 Matches 19; Conservative 3; Mismatches 15; Indels 0; Gaps 0;

OY 1 IVHGDRKKTALGALGALGMIPTLCOAYEYEVHT 37  
 Db 151 LMEGDRKHMIALSTITALGVFTLLQASFEYEAFT 187

RESULT 3  
 ID Q9G7S0: PRELIMINARY; PRT; 267 AA.  
 AC Q9G7S0;  
 DT 01-MAR-2001 (TREMBlrel. 16, Created)  
 DT 01-MAR-2001 (TREMBlrel. 16, Last sequence update)  
 DT 01-JUN-2001 (TREMBlrel. 17, Last annotation update)  
 DE COI1.  
 GN COI1.  
 OS Sus scrofa (Pig).  
 OG Mitochondrion.  
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 OC Mammalia; Eutheria; Cetartiodactyla; Suina; Suidae; Sus.  
 OX NCBI\_TaxID=9823;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RA Kijas J.M.H., Andersson L.;  
 RT "A phylogenetic study of the origin of the domestic pig estimated from  
 the near complete mtDNA genome."  
 RL Submitted (SEP-2000) to the EMBL/GenBank/DBJ databases.  
 DR EMBL; AF304203; AAC28222.1;  
 DR InterPro; IPR000298; CytC\_oxdse\_III.  
 DR Pfam; PF00510; COX3; 1.  
 DR PROSITE; PS50253; COX3; 1.  
 KW Mitochondrion.  
 SQ SEQUENCE 267 AA; 30454 MW; 471AED9C294867E9 CRC64;

Query Match 41.4%; Score 84; DB 8; Length 267;  
 Best Local Similarity 51.4%; Pred. No. 0.0023;  
 Matches 19; Conservative 3; Mismatches 15; Indels 0; Gaps 0;

OY 1 IVHGDRKKTALGALGALGMIPTLCOAYEYEVHT 37  
 Db 151 LMEGDRKHMIALSTITALGVFTLLQASFEYEAFT 187

RESULT 4  
 Q9G290

ID Q9G290: PRELIMINARY; PRT; 278 AA.  
 AC Q9G290;  
 DT 01-MAR-2001 (TREMBlrel. 16, Created)  
 DT 01-MAR-2001 (TREMBlrel. 16, Last sequence update)  
 DE 01-JUN-2001 (TREMBlrel. 17, Last annotation update)  
 GN COI1.  
 OS Sus scrofa (Pig).  
 OG Mitochondrion.  
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 OC Mammalia; Eutheria; Cetartiodactyla; Suina; Suidae; Sus.  
 OX NCBI\_TaxID=9823;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RA Kijas J.M.H., Andersson L.;  
 RT "A phylogenetic study of the origin of the domestic pig estimated from  
 the near complete mtDNA genome."  
 RL Submitted (SEP-2000) to the EMBL/GenBank/DBJ databases.  
 DR EMBL; AF304202; AAC28209.1;  
 DR EMBL; AF304200; AAC28183.1;  
 DR InterPro; IPR000298; CytC\_oxdse\_III.  
 DR Pfam; PF00510; COX3; 1.  
 DR PROSITE; PS50253; COX3; 1.  
 KW Mitochondrion.  
 SQ SEQUENCE 278 AA; 31927 MW; 25A60814452E5A2B CRC64;

Query Match 41.4%; Score 84; DB 8; Length 278;  
 Best Local Similarity 51.4%; Pred. No. 0.0024;  
 Matches 19; Conservative 3; Mismatches 15; Indels 0; Gaps 0;

OY 1 IVHGDRKKTALGALGALGMIPTLCOAYEYEVHT 37  
 Db 151 LMEGDRKHMIALSTITALGVFTLLQASFEYEAFT 187

RESULT 5  
 ID Q9T244: PRELIMINARY; PRT; 305 AA.  
 AC Q9T244;  
 DT 01-MAY-2000 (TREMBlrel. 13, Created)  
 DT 01-MAY-2000 (TREMBlrel. 13, Last sequence update)  
 DT 01-JUN-2001 (TREMBlrel. 17, Last annotation update)  
 DE CYTOCHROME C OXIDASE SUBUNIT 3 (EC 1.9.3.1).  
 GN COX3.  
 OS Phytophthora infestans (Potato late blight fungus).  
 OG Mitochondrion.  
 OC Eukaryota; stramenopiles; Oomycetes; Pythiales; Pythiaceae;  
 OC Phytophthora.  
 OX NCBI\_TaxID=4787;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RC STRAIN=WEST VIRGINIA 4;  
 RA Lang B.F., Forget L.;  
 RT "The mitochondrial genome of phytophthora infestans."  
 RL (In) O'Brien S.J. (eds.);  
 RL Genetic Maps, pp.3:133-135, Cold Spring Harbor Laboratory Press,  
 NY (1992).  
 RN [2]  
 RP SEQUENCE FROM N.A.  
 RC STRAIN=WEST VIRGINIA 4;  
 RA Paquin B., Roewer I., Wang Z., Lang B.F.;  
 RT "A robust fungal phylogeny using the mitochondrially encoded mads  
 protein sequence."  
 RL Can. J. Bot. 73:5180-5185(1995).  
 RN [3]  
 RP SEQUENCE FROM N.A.  
 RC STRAIN=WEST VIRGINIA 4;  
 RA Paquin B., Laforest M.J., Forget L., Roewer I., Wang Z., Longcore J.,  
 RT "The fungal mitochondrial genome project: evolution of fungal  
 mitochondrial genomes and their gene expression.";

RL Curr. Genet. 31:380-395(1997).  
 DR EMBL; U17009; AAF24786.1; -;  
 DR InterPro: IPR000298; CytC\_oxide\_III.  
 DR Pfam: PF00510; COX3; 1.  
 DR PROSITE: PS50253; COX3; 1.  
 DR Oxidoreductase; Mitochondrion.  
 KW SEQUENCE 305 AA; 35262 MW; 72CCE/28F7C7CCC7 CRC64;  
 SQ

Query Match 41.4%; Score 84; DB 8; Length 305;  
 Best Local Similarity 54.5%; Pred. No. 0.0027;  
 Matches 18; Conservative 3; Mismatches 12; Indels 0; Gaps 0;

OY 1 IVHGDRKKTATGATATGAGTTCQAYEYEVHT 37  
 Db 194 IVFGDRKNAITLITLITLAFPSLIQAYEYEVHT 226

RESULT 6

ID 09B629 PRELIMINARY; PRT; 261 AA.

AC 09B629;  
 DT 01-JUN-2001 (TREMBLrel. 17, Created)  
 DT 01-JUN-2001 (TREMBLrel. 17, Last sequence update)  
 DT 01-JUN-2001 (TREMBLrel. 17, Last annotation update)  
 DE CYTOCHROME C OXIDASE SUBUNIT III.  
 GN COIT.  
 OS Anguilla japonica (Japanese eel).  
 OG Mitochondrion.  
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 OC Actinopterygii; Neopterygii; Teleostei; Anguilliformes; Anguillidae;  
 OC Anguillidae; Anguilla.  
 OX NCBI\_TaxID=7937;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RA Inoue J.G., Miya M., Nishida M.;  
 RL Submitted (FEB-2000) to the EMBL/GenBank/DBJ databases.  
 RN [2]  
 RP SEQUENCE FROM N.A.  
 RA Inoue J.G., Miya M., Aoyama J., Ishikawa S., Tsukamoto K., Nishida M.;  
 RT "Complete Mitochondrial DNA Sequence of the Japanese Rel, Anguilla  
 japonica".  
 RT Fisheries Sci. 67:118-125(2001).  
 RL Fishers Sci. 67:118-125(2001).  
 DR EMBL; AB038556; BAB39718.1; -;  
 KW Mitochondrion.  
 SQ SEQUENCE 261 AA; 29717 MW; D6317E3C76DA48A7 CRC64;

Query Match 40.9%; Score 83; DB 8; Length 261;  
 Best Local Similarity 48.6%; Pred. No. 0.0031;  
 Matches 18; Conservative 5; Mismatches 14; Indels 0; Gaps 0;

OY 1 IVHGDRKKTATGATATGAGTTCQAYEYEVHT 37  
 Db 151 IMEGRKQAISLITLITLITLGFYFLQAMEYEAFT 187

RESULT 7

ID 09GAU2 PRELIMINARY; PRT; 260 AA.

AC 09GAU2;  
 DT 01-MAR-2001 (TREMBLrel. 16, Created)  
 DT 01-MAR-2001 (TREMBLrel. 16, Last sequence update)  
 DT 01-JUN-2001 (TREMBLrel. 17, Last annotation update)  
 DE CYTOCHROME C OXIDASE POLYPEPTIDE III.  
 OS Rana sylvatica (Wood Frog).  
 OG Mitochondrion.  
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 OC Amphibia; Batrachia; Anura; Neobatrachia; Ranidae; Rana.  
 OX NCBI\_TaxID=45438;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RC TISSUE=BRAIN;

RA Wu S.-B., Storey K.B.;  
 RT "Up-regulation of mitochondrial ATPase genes under freezing exposure  
 and anoxic treatment by wood frog (Rana sylvatica)".  
 RL Submitted (AUG-1999) to the EMBL/GenBank/DBJ databases.  
 DR EMBL; AF175976; AAG33066.1; -;  
 DR InterPro: IPR000298; CytC\_oxide\_III.  
 DR Pfam: PF00510; COX3; 1.  
 DR PROSITE: PS50253; COX3; 1.  
 DR Mitochondrion.  
 KW SEQUENCE 260 AA; 29565 MW; 5639DC103F2CD186 CRC64;  
 SQ

Query Match 40.4%; Score 82; DB 8; Length 260;  
 Best Local Similarity 51.4%; Pred. No. 0.0042;  
 Matches 19; Conservative 2; Mismatches 16; Indels 0; Gaps 0;

OY 1 IVHGDRKKTATGATATGAGTTCQAYEYEVHT 37  
 Db 151 IMQADRKGTGLALTLITLITLITLGFYFLQAMEYEAFT 187

RESULT 8

ID 09B614 PRELIMINARY; PRT; 261 AA.

AC 09B614;  
 DT 01-JUN-2001 (TREMBLrel. 17, Created)  
 DT 01-JUN-2001 (TREMBLrel. 17, Last sequence update)  
 DT 01-JUN-2001 (TREMBLrel. 17, Last annotation update)  
 DE CYTOCHROME C OXIDASE SUBUNIT III.  
 GN CO III.  
 OS Carassius auratus (Goldfish).  
 OG Mitochondrion.  
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 OC Actinopterygii; Neopterygii; Teleostei; Osteichthyes;  
 OC Cypriniformes; Cyprinidae; Cyprininae; Carassius.  
 OX NCBI\_TaxID=7957;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RA Murakami M.;  
 RL "Carassius auratus cv. Kurodai mitochondrial DNA, complete sequence."  
 DR EMBL; AB045144; BAB40354.1; -;  
 KW Mitochondrion.  
 SQ SEQUENCE 261 AA; 29658 MW; D0DD83AD5BE184FF CRC64;

Query Match 40.4%; Score 82; DB 8; Length 261;  
 Best Local Similarity 48.6%; Pred. No. 0.0042;  
 Matches 18; Conservative 5; Mismatches 14; Indels 0; Gaps 0;

OY 1 IVHGDRKKTATGATATGAGTTCQAYEYEVHT 37  
 Db 151 IMEGRKQAISLITLITLITLGFYFLQAMEYEAFT 187

RESULT 9

ID 09ZY39 PRELIMINARY; PRT; 261 AA.

AC 09ZY39;  
 DT 01-MAY-1999 (TREMBLrel. 10, Created)  
 DT 01-MAY-1999 (TREMBLrel. 10, Last sequence update)  
 DT 01-JUN-2001 (TREMBLrel. 17, Last annotation update)  
 DE CYTOCHROME C OXIDASE SUBUNIT III.  
 GN COX-III.  
 OS Raja radata.  
 OG Mitochondrion.  
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Chondrichthyes;  
 OC Elasmobranchii; Squala; Hypnosquala; Pristiogaster; Batoidae;  
 OC Rajiformes; Rajidae; Raja.  
 OX NCBI\_TaxID=27589;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RC TISSUE=LIVER;

RX MEDLINE=99162577; Pubmed=10051614;  
 RA Rasmussen A.-S.; Arnason U.;  
 RT "Molecular studies suggest that cartilaginous fishes have a terminal  
 position in the piscine tree."  
 RN Proc. Natl. Acad. Sci. U.S.A. 96:2177-2182(1999).  
 [2]  
 RP SEQUENCE FROM N.A.  
 RC TISSUE=LIVER;  
 RA Rasmussen A.-S.;  
 RL Submitted (NOV-1998) to the EMBL/GenBank/DBJ databases.  
 DR EMBL: AF106038; AAD19935.1; -  
 DR HSSP: P18402; 1FPT;  
 DR InterPro: IPR000298; Cytc\_oxdse\_III.  
 DR Pfam: PF00510; COX3; 1.  
 DR PROSITE: PS50253; COX3; 1.  
 KW Mitochondrion.  
 SQ SEQUENCE 261 AA; 29600 MW; 6A9B1A6DC779F788 CRC64;

Query Match 39.9%; Score 81; DB 8; Length 261;  
 Best Local Similarity 45.9%; Pred. No. 0.0057;  
 Matches 17; Conservative 5; Mismatches 15; Indels 0; Gaps 0;  
 OY 1 IVHGRRKTAIGLAIAGLWIFLCOAYEYEVHT 37  
 Db 151 IMEGRRKQVQALTLTLGLGFYFLQAMEYEAAPT 187

RESULT 10  
 O9XN34 PRELIMINARY; PRT; 261 AA.  
 AC O9XN34;  
 DT 01-NOV-1999 (TReMBLrel. 12, Created)  
 DT 01-NOV-1999 (TReMBLrel. 12, Last sequence update)  
 DT 01-JUN-2001 (TReMBLrel. 17, Last annotation update)  
 DE CYTOCHROME OXIDASE SUBUNIT 3.  
 GN COIII.  
 OS Salvelinus fontinalis (Brook trout) (Brook char).  
 OG Mitochondrion.  
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 OC Actinopterygii; Neopterygii; Teleostei; Euteleostei;  
 OC Protacanthopterygii; Salmoniformes; Salmonidae; Salvelinus.  
 OX NCBI\_TaxID=8038;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RA Doiron S., Blier P.U., Bernatchez L.;  
 RT "A comparative analysis of complete sequence of mitochondrial genome  
 between brook char (Salvelinus fontinalis) and arctic char (S.  
 alpinus)."  
 RL Submitted (MAY-1999) to the EMBL/GenBank/DBJ databases.  
 DR EMBL: AF154850; AAD41378.1; -  
 DR HSSP: P18402; 1FPT;  
 DR InterPro: IPR000298; Cytc\_oxdse\_III.  
 DR Pfam: PF00510; COX3; 1.  
 DR PROSITE: PS50253; COX3; 1.  
 KW Mitochondrion.  
 SQ SEQUENCE 261 AA; 29721 MW; 9EAAAC788FE00316 CRC64;

Query Match 39.9%; Score 81; DB 8; Length 261;  
 Best Local Similarity 45.9%; Pred. No. 0.0057;  
 Matches 17; Conservative 5; Mismatches 15; Indels 0; Gaps 0;  
 OY 1 IVHGRRKTAIGLAIAGLWIFLCOAYEYEVHT 37  
 Db 151 IMEGRRKQVQALTLTLGLGFYFLQAMEYEAAPT 187

RESULT 11  
 O9XN26 PRELIMINARY; PRT; 261 AA.  
 AC O9XN26;  
 DT 01-NOV-1999 (TReMBLrel. 12, Created)

DT 01-NOV-1999 (TReMBLrel. 12, Last sequence update)  
 DT 01-JUN-2001 (TReMBLrel. 17, Last annotation update)  
 DE CYTOCHROME OXIDASE SUBUNIT 3.  
 GN COIII.  
 OS Salvelinus alpinus (Arctic char).  
 OG Mitochondrion.  
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 OC Actinopterygii; Neopterygii; Teleostei; Euteleostei;  
 OC Protacanthopterygii; Salmoniformes; Salmonidae; Salvelinus.  
 OX NCBI\_TaxID=8036;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RA Doiron S., Blier P.U., Bernatchez L.;  
 RT "A comparative analysis of complete sequence of mitochondrial genome  
 between brook char (Salvelinus fontinalis) and arctic char (S.  
 alpinus)."  
 RL Submitted (MAY-1999) to the EMBL/GenBank/DBJ databases.  
 DR EMBL: AF154851; AAD41391.1; -  
 DR HSSP: P18402; 1FPT;  
 DR InterPro: IPR000298; Cytc\_oxdse\_III.  
 DR Pfam: PF00510; COX3; 1.  
 DR PROSITE: PS50253; COX3; 1.  
 KW Mitochondrion.  
 SQ SEQUENCE 261 AA; 29751 MW; A96564CE40C15BA9 CRC64;

Query Match 39.9%; Score 81; DB 8; Length 261;  
 Best Local Similarity 45.9%; Pred. No. 0.0057;  
 Matches 17; Conservative 5; Mismatches 15; Indels 0; Gaps 0;  
 OY 1 IVHGRRKTAIGLAIAGLWIFLCOAYEYEVHT 37  
 Db 151 IMEGRRKQVQALTLTLGLGFYFLQAMEYEAAPT 187

RESULT 12  
 O9G607 PRELIMINARY; PRT; 261 AA.  
 AC O9G607;  
 DT 01-MAR-2001 (TReMBLrel. 16, Created)  
 DT 01-MAR-2001 (TReMBLrel. 16, Last sequence update)  
 DT 01-JUN-2001 (TReMBLrel. 17, Last annotation update)  
 DE CYTOCHROME C OXIDASE SUBUNIT III.  
 GN COIIT.  
 OS Diplodus taenia.  
 OG Mitochondrion.  
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 OC Actinopterygii; Neopterygii; Teleostei; Euteleostei; Neoteleostei;  
 OC Stomiiformes; Gonostomatidae; Diplodus.  
 OX NCBI\_TaxID=91975;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RA Miya M., Nishida M.;  
 RT "Setting a new stage for the teleostean molecular systematics: A  
 mitochondrial approach using long PCR and versatile primers."  
 RL Submitted (NOV-1999) to the EMBL/GenBank/DBJ databases.  
 DR EMBL: AB034825; BAB20726.1; -  
 DR InterPro: IPR000298; Cytc\_oxdse\_III.  
 DR Pfam: PF00510; COX3; 1.  
 DR PROSITE: PS50253; COX3; 1.  
 KW Mitochondrion.  
 SQ SEQUENCE 261 AA; 29585 MW; 694856C933298B77 CRC64;

Query Match 39.9%; Score 81; DB 8; Length 261;  
 Best Local Similarity 48.6%; Pred. No. 0.0057;  
 Matches 18; Conservative 5; Mismatches 14; Indels 0; Gaps 0;  
 OY 1 IVHGRRKTAIGLAIAGLWIFLCOAYEYEVHT 37  
 Db 151 IMEGRRKQVQALTLTLGLGFYFLQALEYEAAPT 187



## RESULT 13

09G387 PRELIMINARY; PRT: 261 AA.  
 AC 09G387;  
 DT 01-MAR-2001 (TREMBLrel. 16, Created)  
 DT 01-MAR-2001 (TREMBLrel. 16, last sequence update)  
 DT 01-JUN-2001 (TREMBLrel. 17, last annotation update)  
 DE CYTOCHROME OXIDASE SUBUNIT III.  
 GN COII.  
 OS Echinos telfairi (Lesser hedgehog tenrec).  
 OG Mitochondrion.  
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 OC Mammalia; Eutheria; Insectivora; Tenrecidae; Echinos.  
 OX NCBI\_TaxID=9371;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RA Mouchaty S.K., Gullberg A., Janke A., Arnason U.;  
 RT "Phylogenetic position of the tenrecs (Mammalia: Tenrecidae) of  
 RT Madagascar based on analysis of the complete mitochondrial genome  
 RT sequence of Echinos telfairi.";  
 RL Zool. Scr. 29:307-317(2000).  
 DR EMBL: AJ400734; CAC19398.1; -  
 DR InterPro: IPR000298; CytC\_oxdse\_III.  
 DR Pfam: PF00510; COX3; 1.  
 DR PROSITE: PS50253; COX3; 1.  
 KW Mitochondrion.  
 SQ SEQUENCE 261 AA; 29773 MW; B66DEB567687DD58 CRC64;

## Query Match

39.9%; Score 81; DB 8; Length 261;  
 Best Local Similarity 48.6%; Pred. No. 0.0057;  
 Matches 18; Conservative 5; Mismatches 14; Indels 0; Gaps 0;

OY 1 IVHGRRKTAIGLAIAGLWIFLLQAYEYEVHT 37  
 Db 151 LMEGNRKMIQALSTITALGLYFTLLQAMEYEAFT 187

## RESULT 14

09B263 PRELIMINARY; PRT: 261 AA.  
 AC 09B263;  
 DT 01-JUN-2001 (TREMBLrel. 17, Created)  
 DT 01-JUN-2001 (TREMBLrel. 17, last sequence update)  
 DT 01-JUN-2001 (TREMBLrel. 17, last annotation update)  
 DE CYTOCHROME OXIDASE SUBUNIT III.  
 GN COIII.  
 OS Isodon macrourus (Short-nosed bandicoot).  
 OG Mitochondrion.  
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 OC Mammalia; Metatheria; Peramelomorpha; Peramelidae; Isodon.  
 OX NCBI\_TaxID=37698;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RA Phillips M.J., Lin Y.-H., Harrison G.L., Penny D.;  
 RT "Mitochondrial Genomes of a Bandicoot and a Brush-tail Possum Confirm  
 RT the Monophyly of Australidelphian Marsupials.";  
 RL Proc. R. Soc. Lond. B, Biol. Sci. 0:0-0(2001).  
 RN [2]  
 RP SEQUENCE FROM N.A.  
 RA Lin Y.-H., Phillips M.J.;  
 RL Submitted (MAR-2001) to the EMBL/GenBank/DBJ databases.  
 DR EMBL: AF358864; AAK38287.1; -  
 KW Mitochondrion.  
 SQ SEQUENCE 261 AA; 29966 MW; A108114486943414 CRC64;

## Query Match

39.9%; Score 81; DB 8; Length 261;  
 Best Local Similarity 45.9%; Pred. No. 0.0057;  
 Matches 17; Conservative 6; Mismatches 14; Indels 0; Gaps 0;

OY 1 IVHGRRKTAIGLAIAGLWIFLLQAYEYEVHT 37  
 Db 151 LMEGNRKMIQALSTITALGLYFTLLQAMEYEAFT 187

Db 151 LMEGNRKMIQALSTITALGLYFTLLQAMEYEAFT 187

## RESULT 15

003378 PRELIMINARY; PRT: 261 AA.  
 AC 003378;  
 DT 01-JUL-1997 (TREMBLrel. 04, Created)  
 DT 01-JUL-1997 (TREMBLrel. 04, last sequence update)  
 DT 01-JUN-2001 (TREMBLrel. 17, last annotation update)  
 DE CYTOCHROME OXIDASE SUBUNIT 3.  
 OS Oncorhynchus keta (Chum salmon).  
 OG Mitochondrion.  
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 OC Actinopterygii; Neopterygii; Teleostei; Euteleostei;  
 OC Proteocephali; Salmoniformes; Salmonidae; Oncorhynchus.  
 OX NCBI\_TaxID=8018;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RA STRAIN-TOKORO RIVER, HOKKAIDO, JAPAN;  
 RC MEDLINE=97159563; PubMed=9007022;  
 RA Oohara I., Sawano K., Okazaki T.;  
 RT "Mitochondrial DNA sequence analysis of the masu salmon--phylogeny in  
 RT the genus Oncorhynchus.";  
 RL Mol. Phylogenet. Evol. 7:71-78(1997).  
 DR EMBL: D84147; BAA20154.1; -  
 DR HSSP: P16402; IFFT  
 DR InterPro: IPR000298; CytC\_oxdse\_III.  
 DR Pfam: PF00510; COX3; 1.  
 DR PROSITE: PS50253; COX3; 1.  
 KW Mitochondrion.  
 SQ SEQUENCE 261 AA; 29691 MW; 4EF55462ACD66DF CRC64;

## Query Match

39.4%; Score 80; DB 8; Length 261;  
 Best Local Similarity 45.9%; Pred. No. 0.0077;  
 Matches 17; Conservative 5; Mismatches 15; Indels 0; Gaps 0;

OY 1 IVHGRRKTAIGLAIAGLWIFLLQAYEYEVHT 37  
 Db 151 LMEGNRKMIQALSTITALGLYFTLLQAMEYEAFT 187

Search completed: January 10, 2002, 02:04:05  
 Job time: 217 sec







us-09-712-768-6.rn

ATTORNEY/AGENT INFORMATION:  
NAME: Campbell, Cathryn A.  
REGISTRATION NUMBER: 31,815  
REFERENCE/DOCKET NUMBER: P-AG 9504  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: (619) 535-9001  
TELEFAX: (619) 355-8949  
INFORMATION FOR SEO ID NO.: 5:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 856 base pairs  
TYPE: nucleic acid  
STRANDEDNESS: double  
TOPOLOGY: linear

```
alignment_scores:      Length: 33
      Quality:          Gaps: 0
      Ratio: 2.583
Percent Similarity: 72.727 Percent Identity: 42.424
```

```
alignment_block:
us-09-712-768-6 x us-08-451-096-3 ..
```

Align seg 1/1 to: US-08-451-096-3 from: 1 to: 856

1 lIeVaHISGIYAAsparGlySLySThrAlaIleGlyLeuAlaIleAlI 17  
 521 CFNATGAAACAAACCGAACCCAAATTAATTCAGAGCAGCTTATTCACAT 570  
 17 eGyIeuGLyTTPHlePheThrLeuCySGlnAlaTyGlyTyrrGlu 33  
 571 TTTACTGGGCTCTATTATTCACCTCCATAACAGCCCTCAAGATCTTCAG 619  
 TTTACTGGGCTCTATTATTCACCTCCATAACAGCCCTCAAGATCTTCAG 619

```
seq_name: /cgn2_6/ptodata/2/ina/5B_COMB.seq:US-08-810-599-82
```

```
seq_documentation_block:
  sequence 82, Application US/08810599
```

GENERAL INFORMATION: PARKER, W. Davis  
APPLICANT: HEHRNSTADT, Corinna  
APPLICANT: GHOSH, Soumitra S.  
APPLICANT: FAY, Eoin  
TITLE OF INVENTION: Methods for Detecting Mitochondrial Mutations  
TITLE OF INVENTION: Diagnostic for Alzheimer's Disease and Methods for Determining  
TITLE OF INVENTION: of Mitochondrial Nucleic Acid  
NUMBER OF SEQUENCES: 82  
CORRESPONDENCE ADDRESS: Kenyon  
ADDRESSEE: Kenyon & Kenyon  
N.W., Suite 600  
Avenue

1 CORRESPONDENCE: Kenyon & Kenyon  
2  
3 ADDRESS: 1025 Connecticut Avenue, N.W., Suite 600  
4 STREET: Washington  
5 CITY: Washington  
6 STATE: D.C.  
7 COUNTRY: US  
8 ZIP: 20036  
9  
10 COMPUTER READABLE FORM:  
11 MEDIUM TYPE: 3.25" Floppy disk  
12  
13 COMPUTER: IBM PC compatible  
14 OPERATING SYSTEM: PC-DOS/MS-DOS  
15 SOFTWARE: Wordperfect 6.1 for Windows  
16  
17 CURRENT APPLICATION DATA:  
18 APPLICATION NUMBER: US/08/810.599  
19  
20 FILING DATE: Concurrent Herewith  
21  
22 CLASSIFICATION: 436  
23  
24 PRIOR APPLICATION DATA:  
25 APPLICATION NUMBER: US 08/757,438  
26 FILING DATE: 27 NO. 5976798 1996  
27 APPLICATION NUMBER: US 08/614,072  
28  
29 FILING DATE: 12 Mar 1996  
30 APPLICATION NUMBER: US 08/536,036  
31  
32 FILING DATE: 29 Sep 1995  
33 APPLICATION NUMBER: US 08/414,969  
34  
35 FILING DATE: 31 Mar 1995

APPLICATION NUMBER: US 08/413,740  
FILING DATE: 30 Mar 1995  
APPLICATION NUMBER: US 08/410,658  
FILING DATE: 24 MARCH 1995  
APPLICATION NUMBER: US 08/397,808  
FILING DATE: 3 Mar 1995  
APPLICATION NUMBER: US 08/219,843  
FILING DATE: 30 MARCH 1994  
ATTORNEY/AGENT INFORMATION:  
NAME: Toifenetti, Judith L.  
REGISTRATION NUMBER: 39,048  
REFERENCE/DOCKET NUMBER: 2105/17  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: 202-429-1776  
TELEFAX: 202-429-0796  
INFORMATION FOR SEQ ID NO: 82:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 926 base pairs  
TYPE: nucleic acid  
STRANDEDNESS: double  
TOPOLOGY: linear  
MOLECULE TYPE: other nucleic acid  
HYPOTHETICAL: NO  
ANTI-SENSE: NO  
US-08-810-599-82

```
alignment_scores: 33
  quality: 62.00
  ratio: 2.583
  percent_similarity: 72.727
  percent_identity: 42.424
```

alignment\_block:  
US-08-810-599-82

Align seg 1/1 to: US-08-810-599-82 from: 1 to: 926

[illegible]

```
seq_name: /cgn2_6/ptodata/2/ina/bb_COMB.seq.00 00 00
```

```
seq_documentation_block:
  . Sequence 2, Application US/09097889
```

1 GENERAL INFORMATION:  
2 Applicant: Herrnsrad, Corrina  
3 Applicant: Chosh, Soumitra S.  
4 Applicant: Davis, Robert E.  
5 TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR IDENTIFYING  
6 TITLE OF INVENTION: AGENTS THAT QUANTITATIVELY ALTER DETECTABLE  
7 TITLE OF INVENTION: EXTRAMITOCHONDRIAL DNA RATIOS  
8 TITLE OF INVENTION: MITOCHONDRIAL DNA  
9 NUMBER OF SEQUENCES: 26  
10 CORRESPONDENCE ADDRESS: Harry HRP

CORRESPONDENCE: SEED and BERRY LLP  
ADDRESS: 6300 Columbia Center, 701 Fifth Avenue  
STREET:  
CITY: Seattle  
STATE: Washington  
COUNTRY: USA  
ZIP: 98104  
COMPUTER READABLE FORM:  
MEDIUM TYPE: Floppy disk  
COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: PatentIn Release #1.0, Version #1.30  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/09/097, 889  
FILING DATE: 15-JUN-1998

1

STRANDEDNESS: single

```

ORIGINAL SOURCE:
ORGANISM: PAG1250UP
US-08-998-416-306

alignment_scores:
Quality: 56.00 Length: 31
Ratio: 2.667 Gaps: 0
Percent Similarity: 67.742 Percent Identity: 41.935

alignment_block:
US-09-712-768-6 x US-08-998-416-306 ..
Align seg 1/1 to: US-08-998-416-306 from: 1 to: 847

11levalhslgIyaspArglyslYstHAlalleIleIleuAlalleAla11 17
:::|||||:::|||||:::|||||:::|||||:::|||||:::|||||:::
459TtATATTGAAGTAATAGAAAACATGCTTATTCAGGTTACTTATTACTTT 508
17egIleuGlytrPIlePheThrIleuCySGlnAlaIlyrgIuYr 31
::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|
509CTGATTAAATGTTACATTGTGTTATGTCATCATATTAATGAATAT 551

seq_name: /cgn2_6/ptodata/2/ina/6A_COMB.seq:US-09-356-952-12

seq_documentation_block:
; Sequence 12, Application US/09356952
; Patent No. 6117663
; GENERAL INFORMATION:
; APPLICANT: Horlack-Sjodin, Ann
; APPLICANT: Margaret, S. M.
; APPLICANT: Bor-Sogil, Dafna
; APPLICANT: Cole, Philip
; APPLICANT: Kuriyan, John
; TITLE OF INVENTION: A CRYSTAL OF A RAS-SOS COMPLEX AND METHODS OF USE
; TITLE OF INVENTION: THEREOF
; FILE REFERENCE: 600-1-228N
; CURRENT FILING DATE: 1999-07-19
; EARLIER APPLICATION NUMBER: 60/093,631
; EARLIER FILING DATE: 1998-07-21
; NUMBER OF SEQ ID NOS: 14
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 12
; LENGTH: 43676
; TYPE: DNA
; ORGANISM: Saccharomyces cerevisiae
US-09-356-952-12

alignment_scores:
Quality: 54.00 Length: 32
Ratio: 2.842 Gaps: 1
Percent Similarity: 59.375 Percent Identity: 46.875

alignment_block:
US-09-712-768-6 x US-09-356-952-12/rev ..
Align seg 1/1 to reverse of: US-09-356-952-12 from: 1 to: 43676

11lleGIleuAlalleAlalleIleGIleuGIlytrPIle.....Ph 23
|||::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|
497ATTAGGCTACAGAAACATATAGATTTGGACATCCATCCATAGAACCTCTT 448
23ethIleuCySGlnAlaIlyrGIuYrYrgIuIleValhIsthrGlu 38
|||||:::|||||:::|||||:::|||||:::|||||:::|||||:::
447TACGTTGTGCTGTCTTATTTATGTAATCATCATGTCAGTACGAG 402

seq_name: /cgn2_6/ptodata/2/ina/PCTUS_COMB.seq:PCT-US91-09422-20

seq_documentation_block:
; Sequence 20, Application PC/TUS9109422
; GENERAL INFORMATION:

```

APPLICANT: Mulvihill, Eileen R.  
 APPLICANT: Hagen, Frederick S.  
 APPLICANT: Houamed, Khalel M.  
 APPLICANT: Almers, Wolfhard  
 TITLE OF INVENTION: G PROTEIN-COUPLED GLUTAMATE RECEPTORS  
 NUMBER OF SEQUENCES: 33  
 CORRESPONDENCE ADDRESSES: 33  
 ADDRESS: Townsend and Townsend  
 STREET: One Market Plaza, Stewart Street Tower  
 CITY: San Francisco  
 STATE: California  
 COUNTRY: USA  
 ZIP: 94105-1492  
 COMPUTER READABLE FORM:  
 MEDIUM TYPE: Floppy disk  
 OPERATING SYSTEM: PC-DOS/MS-DOS  
 SOFTWARE: Patent Release #1.0, Version #1.25  
 CURRENT APPLICATION DATA:  
 APPLICATION NUMBER: PCT/US91/09422  
 FILING DATE: 19911212  
 CLASSIFICATION: 435  
 PRIOR APPLICATION DATA:  
 APPLICATION NUMBER: US 07/672,007  
 FILING DATE: 18-MAR-1991  
 PRIOR APPLICATION DATA:  
 APPLICATION NUMBER: US 07/648,481  
 FILING DATE: 30-JAN-1991  
 APPLICATION DATA:  
 FILING DATE: 12-DEC-1990  
 ATTORNEY/AGENT INFORMATION:  
 NAME: Parmelee, Steven W.  
 REGISTRATION NUMBER: 31,990  
 TELECOMMUNICATION INFORMATION:  
 TELEPHONE: (206) 467-9600  
 TELEFAX: (415) 543-5043  
 INFORMATION FOR SEQ ID NO: 20:  
 SEQUENCE CHARACTERISTICS:  
 LENGTH: 2426 base pairs  
 TYPE: NUCLEIC ACID  
 STRANDEDNESS: single  
 TOPOLOGY: linear  
 MOLECULE TYPE: cDNA  
 IMMEDIATE SOURCE:  
 CLONE: SR13  
 PCT-US91-09422-20

## alignment\_scores:

Quality: 50.00 Length: 28  
 Ratio: 2.941 Gaps: 1  
 Percent Similarity: 60.714 Percent Identity: 39.286

## alignment\_block:

US-09-712-768-6 x PCT-US91-09422-20 ..

Align seg 1/1 to: PCT-US91-09422-20 from: 1 to: 2426

4 GlyasparglysthrAlaIleGlyLeuAlaIleAlaIleGlyLeuG1 20

1291 GGGGAGGGAAGAAGACTGTGAGGCGATGCTGCTGC..... 1329

20 YTPrllePheThrLeuGlyGlnAlaIleGlyGluTyr 31

1330 .TGGCACTGGGACCCCTGCACCGGGGTACACAGTAC 1362

seq\_name: /cgn2\_6/ptodata/2/1na/PCTUS\_COMB.seq:PCT-US91-09422-20

seq\_documentation\_block:

Sequence 18, Application PC/TUS9109422

GENERAL INFORMATION:

APPLICANT: Mulvihill, Eileen R.  
 APPLICANT: Hagen, Frederick S.  
 APPLICANT: Houamed, Khalel M.  
 APPLICANT: Almers, Wolfhard  
 TITLE OF INVENTION: G PROTEIN-COUPLED GLUTAMATE RECEPTORS  
 NUMBER OF SEQUENCES: 33  
 CORRESPONDENCE ADDRESSES: 33  
 ADDRESS: Townsend and Townsend  
 STREET: One Market Plaza, Stewart Street Tower  
 CITY: San Francisco  
 STATE: California  
 COUNTRY: USA  
 ZIP: 94105-1492  
 COMPUTER READABLE FORM:  
 MEDIUM TYPE: Floppy disk  
 OPERATING SYSTEM: PC-DOS/MS-DOS  
 SOFTWARE: Patent Release #1.0, Version #1.25  
 CURRENT APPLICATION DATA:  
 APPLICATION NUMBER: PCT/US91/09422  
 FILING DATE: 19911212  
 CLASSIFICATION: 435  
 PRIOR APPLICATION DATA:  
 APPLICATION NUMBER: US 07/672,007  
 FILING DATE: 18-MAR-1991  
 PRIOR APPLICATION DATA:  
 APPLICATION NUMBER: US 07/648,481  
 FILING DATE: 30-JAN-1991  
 APPLICATION DATA:  
 FILING DATE: 12-DEC-1990  
 ATTORNEY/AGENT INFORMATION:  
 NAME: Parmelee, Steven W.  
 REGISTRATION NUMBER: 31,990  
 TELECOMMUNICATION INFORMATION:  
 TELEPHONE: (206) 467-9600  
 TELEFAX: (415) 543-5043  
 INFORMATION FOR SEQ ID NO: 18:  
 SEQUENCE CHARACTERISTICS:  
 LENGTH: 4095 base pairs  
 TYPE: NUCLEIC ACID  
 STRANDEDNESS: single  
 TOPOLOGY: linear  
 MOLECULE TYPE: cDNA  
 IMMEDIATE SOURCE:  
 CLONE: SN30  
 FEATURE:  
 NAME/KEY: CDS  
 LOCATION: 463..3198  
 PCT-US91-09422-18

## alignment\_scores:

Quality: 50.00 Length: 28  
 Ratio: 2.941 Gaps: 1  
 Percent Similarity: 60.714 Percent Identity: 39.286

## alignment\_block:

US-09-712-768-6 x PCT-US91-09422-18 ..

Align seg 1/1 to: PCT-US91-09422-18 from: 1 to: 4095

4 GlyasparglysthrAlaIleGlyLeuAlaIleAlaIleGlyLeuG1 20

2041 GGGGAGGGAAGAAGACTGTGAGGCGATGCTGCTGC..... 2079

20 YTPrllePheThrLeuGlyGlnAlaIleGlyGluTyr 31

2080 .TGGCACTGGGACCCCTGCACCGGGGTACACAGTAC 2112

seq\_name: /cgn2\_6/ptodata/2/1na/5A\_COMB.seq:US-07-689-008-4



```

seq_documentation_block:
; Sequence 4, Application US/07689008
; Patent No. 5268274
; GENERAL INFORMATION:
; APPLICANT: Ben-Bassat, Arle
; APPLICANT: Calhoun, Roger D
; APPLICANT: Fear, Anna L
; APPLICANT: Gelfand, David H
; APPLICANT: Meade, James H
; APPLICANT: Tal, Roni
; APPLICANT: Wong, Hing
; APPLICANT: Benzman, Moshe
; TITLE OF INVENTION: METHODS AND NUCLEIC ACID SEQUENCES FOR THE
; TITLE OF INVENTION: EXPRESSION OF CELLULOSE SYNTHASE OPERON
; NUMBER OF SEQUENCES: 15
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: McCutchen, Doyle, Brown & Eversen
; STREET: Three Embarcadero Center
; CITY: San Francisco
; STATE: California
; COUNTRY: USA
; ZIP: 94111
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.24
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/07/689,008
; FILING DATE: 19910422
; CLASSIFICATION: 435
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 337,194
; FILING DATE: 12-APR-1989
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 496,236
; FILING DATE: 23-MAR-1990
; ATTORNEY/AGENT INFORMATION:
; NAME: Murphy, Lisabeth Felix
; REGISTRATION NUMBER: 31547
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (415) 393-2286
; TELEFAX: (415) 393-2000
; TELEX: 340817 MACPAG SFO
; INFORMATION FOR SEQ ID NO: 4:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 2406 base pairs
; TYPE: NUCLEIC ACID
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: DNA (genomic)
US-07-689-008-4

alignment_scores:
Quality: 49.50 Length: 44
Ratio: 2.062 Gaps: 3
Percent Similarity: 54.545 Percent Identity: 31.818

alignment_block:
US-09-712-768-6 x US-07-689-008-4/rev ..
Align seg 1/1 to reverse of: US-07-689-008-4 from: 1 to: 2406

3 HisGlyAsp.....ArgLysLysThrAlaIleGlyLeuAlaIle.. 15
|||||
2244 CACGCTGCCAACCGATACGACGCGGAGTAGCATGTCAGATCAT 2195
16 .....AlaIleGlyLeuGlyTTPiIlePheThrL 25
|||||
2194 CGCCATGCCGACACACAGAGTCCGCTGAGATCTTGCCGCTTTTACGC 2145
25 eucysglnalatyrglutyrtgltuileval 35

```

```

seq_documentation_block:
; Sequence 1, Application US/09215131
; Patent No. 6030834
; GENERAL INFORMATION:
; APPLICANT: Chu, Keting
; APPLICANT: Pot, David
; TITLE OF INVENTION: IKK Beta Regulates Transcription Factors
; FILE REFERENCE: 1449.002
; CURRENT APPLICATION NUMBER: US/09/215,131
; CURRENT FILING DATE: 1998-12-18
; NUMBER OF SEQ ID NOS: 3
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 1
; LENGTH: 3966
; TYPE: DNA
; ORGANISM: human
US-09-215-131-1

alignment_scores:
Quality: 49.50 Length: 30
Ratio: 2.605 Gaps: 2
Percent Similarity: 63.333 Percent Identity: 36.667

alignment_block:
US-09-712-768-6 x US-09-215-131-1 ..
Align seg 1/1 to: US-09-215-131-1 from: 1 to: 3966

3 HisGlyAspArgLysLysThrAlaIle.....GlyLe 13
|||||
3483 CATGAGATGAGATGAGATGAGATCTGTGTCAGTAGGAGGAGCTAT 3532
13 uAlaIleAlaIleGlyLeuGlyTTPiIlePheThrLeuGly 26
: : : : :
3533 TCAGTTTCACTTGTCTTTG...TGGAGATTCACACTATGC 3569

seq_name: /cgn2_6/ptodata/2/ina/6A_COMB.seq:US-09-222-734-1

seq_documentation_block:
; Sequence 1, Application US/09222734A
; Patent No. 6077701
; GENERAL INFORMATION:
; APPLICANT: Chu, Keting
; APPLICANT: Pot, David
; TITLE OF INVENTION: IKK-beta Regulates Transcription Factors
; FILE REFERENCE: 12441.78080
; CURRENT APPLICATION NUMBER: US/09/222,734A
; CURRENT FILING DATE: 1998-12-29
; EARLIER APPLICATION NUMBER: 09/215,131
; EARLIER FILING DATE: 1998-12-18
; EARLIER APPLICATION NUMBER: 60/068,954
; EARLIER FILING DATE: 1997-12-30
; NUMBER OF SEQ ID NOS: 3
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 1
; LENGTH: 3966
; TYPE: DNA
; ORGANISM: Homo sapiens
US-09-222-734-1

alignment_scores:
Quality: 49.50 Length: 30
Ratio: 2.605 Gaps: 2
Percent Similarity: 63.333 Percent Identity: 36.667

alignment_block:

```



GenCore version 4.5  
Copyright (c) 1993 - 2000 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: January 10, 2002, 00:03:47 ; Search time 1538.44 seconds  
(without alignments)  
796.273 Million cell updates/sec

Title: US-09-712-768-5  
Perfect score: 1 atgcgtccagcgagccgca.....atgaatcgtccatccgca 114  
Sequence:

Scoring table: IDENTITY\_NUC  
Gapop 10.0 , Gapext 1.0

Searched: 11351937 seqs, 537289281 residues

Total number of hits satisfying chosen parameters: 22703874

Minimum DB seq length: 0  
Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%  
Maximum Match 100%

Listing first 45 summaries

Database :

EST: \*  
1: em\_estfun: \*  
2: em\_esthm: \*  
3: em\_estin: \*  
4: em\_estom: \*  
5: em\_estpl: \*  
6: em\_estba: \*  
7: em\_estro: \*  
8: em\_estov: \*  
9: em\_hlc: \*  
10: qb\_est1: \*  
11: qb\_est2: \*  
12: qb\_hlc: \*  
13: qb\_gss: \*  
14: em\_gss\_fun: \*  
15: em\_gss\_hum: \*  
16: em\_gss\_inv: \*  
17: em\_gss\_pln: \*  
18: em\_gss\_pro: \*  
19: em\_gss\_pro: \*  
20: em\_gss\_vrt: \*  
21: em\_gss\_other: \*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

# SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	35.4	31.1	435	11 C82943	C82943 C82943 rabb
2	35.4	30.2	435	11 C83799	C83799 C83799 rabb
3	34.4	30.2	436	11 BF890606	BF890606 CM1-MT010
4	34.4	30.2	408	10 AM619231	AM619231 583 MARC
5	34.4	30.2	454	11 B1185787	B1185787 UNL-P-FN-
6	34.4	30.2	474	10 AM507072	AM507072 SOF Band
7	34.4	30.2	482	11 B1337977	B1337977 361613 MA
8	34.4	30.2	485	10 AM619917	AM619917 8001 MARC
9	34.4	30.2	486	11 B1337959	B1337959 361590 MA
10	34.4	30.2	521	11 B1340541	B1340541 365841 MA
11	34.4	30.2	530	11 BF080960	BF080960 233775 MA
12	34.4	30.2	539	10 BE231856	BE231856 136697 MA

13	34.4	30.2	590	10 AM619533	AM619533 7599 MARC
14	34.4	30.2	618	10 AM619953	AM619953 8039 MARC
15	34.4	30.2	647	10 AM619624	AM619624 7692 MARC
16	34.4	30.2	781	11 298830	298830 298830 p17T
17	33.8	29.6	297	11 BF366213	BF366213 112-MT010
18	33.8	29.6	837	11 BF43866	BF43866 602722621
19	33.6	29.5	425	10 AM260970	AM260970 EQUK0167
20	33.3	28.9	561	10 AV590980	AV590980 AV590980
21	32.8	28.8	397	10 BE032165	BE032165 131006 MA
22	32.6	28.6	556	10 AV662541	AV662541 AV662541
23	32.2	28.2	390	11 BF336164	BF336164 OVA-CT049
24	31.4	27.5	503	11 BE518508	BE518508 EST00084
25	31.2	27.4	531	11 BF430098	BF430098 256049 MA
26	31.2	27.2	302	10 AV665359	AV665359 AV665359
27	31.1	27.2	302	10 AM414509	AM414509 48216 MAR
28	31.1	27.2	318	11 BF430098	BF430098 352745 MA
29	31.1	27.2	354	11 BG834894	BG834894 353401 MA
30	31.1	27.2	438	11 BG834847	BG834847 353209 MA
31	31.1	27.2	543	10 AV665622	AV665622 AV665622
32	31.1	27.2	543	10 AM480914	AM480914 33560 MAR
33	30.6	26.8	682	10 BF612450	BF612450 dae18c05
34	30.6	26.7	608	10 AV662352	AV662352 AV662352
35	30.4	26.7	644	13 AQ957803	AQ957803 LERAR76TR
36	30.4	26.7	769	13 AQ957802	AQ957802 LERAR76TR
37	30.4	26.5	431	10 AV590979	AV590979 AV590979
38	30.2	26.5	453	10 AV589848	AV589848 AV589848
39	30.2	26.5	532	10 AV588658	AV588658 AV588658
40	30.2	26.5	535	11 BG555571	BG555571 dF0507.X
41	30.2	26.5	550	10 AV591517	AV591517 AV591517
42	30.2	26.5	550	10 AV664015	AV664015 AV664015
43	30.2	26.5	558	10 AV590621	AV590621 AV590621
44	30.2	26.5	558	10 AV589292	AV589292 AV589292
45	30.2	26.5	564	10 AV589292	AV589292 AV589292

## ALIGNMENTS

RESULT 1	C82943	435 bp	mRNA	EST	26-MAR-1999
LOCUS	C82943				
DEFINITION	C82943		rabbit corneal endothelial cell	Oryctolagus cuniculus	CDNA
ACCESSION	C82943.1	GI:3061900			
VERSION	C82943.1				
KEYWORDS	EST.				
SOURCE	EST.				
ORGANISM	Oryctolagus cuniculus				
	rabbit.				
REFERENCE	Fujimaki, T., Hotta, Y., Sakuma, H., Fujiki, K. and Kanai, A.				
AUTHORS	1 (bases 1 to 435)				
TITLE	Large-scale sequencing of the rabbit corneal endothelial cDNA				
JOURNAL	Cornea 18 (1), 109-114 (1999)				
MEDLINE	99110237				
COMMENT	Contact: Takuro Fujimaki				
	Department of Ophthalmology				
	Yuntendo University School of Medicine				
	3-1-3, Hongo, Bunkyo-ku, Tokyo, Japan, Tokyo 113, Japan				
	Tel: +81-3-5802-1092				
	Fax: +81-3-3817-0260				
	Email: fujimaki@med.yuntendo.ac.jp.				
FEATURES	Location/Qualifiers				
source	1..435				
	/organism="Oryctolagus cuniculus"				
	/db_xref="taxon:9986"				
	/clone="1414"				
	/clone_1ib="rabbit corneal endothelial cell"				
	/cell_type="endothelial cell"				
	/note="Organ: cornea"				
BASE COUNT	109 a	120 c	74 g	132 t	
ORIGIN					

VERSION  
BE890606.1  
GT:12282065  
EST.  
EXWORDS  
DURCE  
human.

	RESULT	4
LOCUS	AM619231	
DEFINITION	AM619231	408 bp
ACCESSION	583 MARC	mRNA
VERSION	AM619231	EST
KEYWORDS	AM619231.1	24-MAR-2000
SOURCE	EST.	
ORGANISM	PI9.	
	<i>Sus scrofa</i>	
	Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;	
	Mammalia; Eutheria; Cetartiodactyla; Suidae; Suidae; Sus.	
REFERENCE	Smith,T.P.L., Pahrenkrug,S.C., Rohrer,G.A., Slimmen,F.A., Rexroad	
AUTHORS	,C.E. and Keele,J.W.	

TITLE	Mapping of expressed sequence tags from a porcine early embryonic
JOURNAL	Anim. Genet. 32 (2), 66-72 (2001)
MEDLINE	21314990
COMMENT	Contact: Smith TPL usda ars us Meat Animal Research Center

FEATURES  
source

```

SOURCE
/organism="Sus scrofa"
/db_xref="taxon:9823"
/clone_1fb="MARC_PBE"
/tissue_type="day 12 whole embryos"
/lab_host="XLOLR"
/note="Vector: pBLUESCRIPT SK-; Site_1: EcoRI; Site_2:
XhoI; Library made from pool of embryos in spherical and
filamentous stages of development (7.58 and 92.58",
respectively, of each stage) as described in Choi, et al,
Endocrinology 137, 1457-67, 1996."
BASE COUNT      114 a      125 c      68 g      101 t
ORIGIN

```

Query Match	30.2%	Score 34.4;	DB 10;	Length 408;
Best Local Similarity	60.9%	Pred. No. 0.3;		
56. Conservative matches	0;	Mismatches	36;	Indels 0; Gaps 0

[illegible]

tail. The sequence tag present in the cDNA between the NotI site and the oligo-dT track served to verify it as a clone from the normalized porcine ovarian follicles library  
Seq primer: M13 -29  
POLYA=Yes.

BASE COUNT	111 a	70 c	132 g	141 c
ORIGIN				

Query Match	30.2%;	Score 34.4;	DB 11;	Length 454;
Best Local Similarity	60.9%;	Pred. No. 0.31;		
Matches	56;	Conservative	0;	Mismatches 36;
				Indels 0;
				Gaps 0.

RESULT	6
AM507072/c	
LOCUS	
DEFINITION	
ACCESSION	
VERSION	
KEYWORDS	
SOURCE	
ORGANISM	
REFERENCE	
AUTHORS	
TITLE	
JOURNAL	
COMMENT	

AM507072 474 bp mRNA EST 13-MAR-2000  
 SOF Band 83 pig ovarian follicle cDNA Sus scrofa CDNA 3' similar to  
 Cytochrome c oxidase subunit 3, mRNA sequence.  
 AM507072  
 AM507072.1 GI:7145591  
 EST.  
 Pig.  
 Sus. scrofa  
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 Mammalia; Eutheria; Cetartiodactyla; Suidae; Sus.  
 1 (bases 1 to 474)  
 Gladney,C.D., Johnson,R.K. and Pomp,D.  
 Evaluation of gene expression in ovarian follicles of pigs selected  
 for reproduction  
 unpublished (2000)  
 Contact: Christy Gladney  
 Molecular Genetics Lab

Email: dponm@unl.edu  
The sequence contained an oligo-dT track that was present in the oligonucleotide that was used to prime the synthesis of first strand cDNA and therefore this may represent a bonafide poly A





## PLK PRIMERS

FORWARD: AGGAACACGATATGACCAT  
 BACKWARD: GTTTCGCCAGTCACGACG  
 Plate: 53      row: M      column: 16  
 Seq primer: ATTTAGCTGACACTATAG  
 Location/Owner: 4444

1. .530" walllets

```

/organism="Sus scrofa"
/db_xref="taxon:9823"
/clone_lib="MARC 2Pig"
/tissue_type="pooled"
/lab_host="DH10B"
/notes="Vector: pCMV SPORT6; Site:1: XbaI; Site:2: XhoI"
library made from pooled tissue from testis, ovary,
endometrium, hypothalamus, pituitary, and placenta."
146 a 85 c 142 t

```

Query Match	30.28;	Score 34.4;	DB 11;	Length 530;
Best Local Similarity	60.98;	Pred. No. 0.32;		
Matches	56;	Conservative	0;	Mismatches 15

[illegible]

RESULT	12
BE231856	
LOCUS	BE231856
DEFINITION	559 bp mRNA
ACCESSION	136697 MARC
REVISION	BE231856
KEYWORDS	PIG Sus scrofa CDNA 5' mRNA sequence.
PROBES	m33307 .
DATE	10-JUL-2000

KEYWORDS	EST.	GI: 90165/4
SOURCE	pig.	
ORGANISM	Sus scrofa	

**REFERENCE**  
**AUTHORS**

mammalia; Eutheria; Cetartiodactyla; Euteleostomi,  
1 (bases 1 to 559)  
Fahrenkrug S C  
Eutheria;  
Sauria; Vertebrata; Euteleostomi,  
Sus.

TITLE	Author(s)
Design and Use of two ...	Laegreid, W.W.

JOURNAL  
COMMENT  
Unpublished (2000)  
Contact: Smith TP.

PO Box 166, University of Maryland System  
 PO Box 166, Maryland Center for Health and Human Sciences  
 Tel.: 402 762 4396  
 Fax: 402 762 4390  
 Email: smith@email.marc.usda.gov  
 Single pass sequencing bases called and alt-trimmed with phred  
 v0.806094.e. Vector identified by cross-match with the -mnscore 1E  
 and -mismatch 12 options.  
 CCR Primers

```

FEATURES
  source      Location/Qualifiers
1..559

```

```

/organism="Sus scrofa"
/db_xref="taxon:9823"
/clone_lib="MARC 1P1c"
/tissue_type="pooled"
/lab_host="DH10B"
/note="Vector: pCMV SPORT6; Site_1: XbaI; Site_2: XhoI"

```

BASE COUNT	ORIGIN
161 a	Library made from pooled tissue from day 11, 13, 15, 20, and 30 embryos."
159 c	
91 g	
148 t	

Query Match	30.2%	Score 34.4;	DB 10;	Length 559;
Best Local Similarity	60.9%;	Pred. No. 0.32;		
Matches	56;	Conservative	0;	Mismatches 30

	8	10	12	14	16	18	20	22	24	26	28	30	32	34	36	38	40	42	44	46	48	50	52	54	56	58	60	62	64	66	68	70	72	74	76	78	80	82	84	86	88	90	92	94	96	98	100	102	104	106	108	110	112	114	116	118	120	122	124	126	128	130	132	134	136	138	140	142	144	146	148	150	152	154	156	158	160	162	164	166	168	170	172	174	176	178	180	182	184	186	188	190	192	194	196	198	200	202	204	206	208	210	212	214	216	218	220	222	224	226	228	230	232	234	236	238	240	242	244	246	248	250	252	254	256	258	260	262	264	266	268	270	272	274	276	278	280	282	284	286	288	290	292	294	296	298	300	302	304	306	308	310	312	314	316	318	320	322	324	326	328	330	332	334	336	338	340	342	344	346	348	350	352	354	356	358	360	362	364	366	368	370	372	374	376	378	380	382	384	386	388	390	392	394	396	398	400	402	404	406	408	410	412	414	416	418	420	422	424	426	428	430	432	434	436	438	440	442	444	446	448	450	452	454	456	458	460	462	464	466	468	470	472	474	476	478	480	482	484	486	488	490	492	494	496	498	500	502	504	506	508	510	512	514	516	518	520	522	524	526	528	530	532	534	536	538	540	542	544	546	548	550	552	554	556	558	560	562	564	566	568	570	572	574	576	578	580	582	584	586	588	590	592	594	596	598	600	602	604	606	608	610	612	614	616	618	620	622	624	626	628	630	632	634	636	638	640	642	644	646	648	650	652	654	656	658	660	662	664	666	668	670	672	674	676	678	680	682	684	686	688	690	692	694	696	698	700	702	704	706	708	710	712	714	716	718	720	722	724	726	728	730	732	734	736	738	740	742	744	746	748	750	752	754	756	758	760	762	764	766	768	770	772	774	776	778	780	782	784	786	788	790	792	794	796	798	800	802	804	806	808	810	812	814	816	818	820	822	824	826	828	830	832	834	836	838	840	842	844	846	848	850	852	854	856	858	860	862	864	866	868	870	872	874	876	878	880	882	884	886	888	890	892	894	896	898	900	902	904	906	908	910	912	914	916	918	920	922	92
--	---	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	----

	RESULT	13
AW619533		
LOCUS	AW619533	590 bp
DEFINITION	7599_MARC_PBE Sus scrofa CDNA 5'	mRNA EST
ACCESSION	AW619533	
VERSION	AW619533.1	mRNA sequence.
		24-MAR-2000

ORGANISM *Sus scrofa*

REFERENCE  
AUTHORS  
1 (bases 1 to 590)  
Smith, T. P. L., Fabrenbrun G. C. 1971. *Journal of Molecular Biology* 59: 1-590.

TITLE	C.E. and Keele, J.W.: Mapping of expressed sequence tags from a porcine early embryonic cDNA library
JOURNAL	Anim. Genet. 32 (2), 66-72 (2001)
MEDLINE	21314990
COMMENT	Contact: Smith TP.

USDA, ARS, US Meat Animal Research Center  
PO Box 166, Clay Center, NE 68933-0166, USA  
Tel.: 402 762 4366  
Fax: 402 762 4390  
Email: smithem@ma.ars.usda.gov  
Single-pass sequencing. Bases called and alt-trimmed with phred  
v0.980904. Vectors identified by cross-match with the minuscore  
and -mismatch 12 options.  
PCR primers

FORWARD: GGAACAGCTATGACCATG  
 BACKWARD: GTAAACGACGGCCAGT  
 Seq primer: AATTAACCTTCACCTAAGGG

Source

Location/Qualifiers
1. .590

```

/olganism="Sus scrofa"
/db_xref="taxon:9823"
/clone_lib="MARC PBE"
/issue_type="Day 12 whole embryos"
/lab_host="X10LR"
/notes="Vector: pBluescript SR-; Site_1: EcoRI; Site_2:
XhoI; Library made from pool of embryos in spherical and
filamentous stages of development (7.5% and 92.5%,
respectively, of each stage) as described in Choi et al.,
Endocrinology 137, 1457-66, 1996."
BASE COUNT
ORIGIN
173 a
171 c      89 g      157 t

```

Query Match	30.28;	Score 34.4;	DB 10;	Length 590;
Best Local Similarity	60.9%;	Pred. NO. 0.33;		
Matches 56;	Conservative 0;	Mismatches 36;	Indels 0;	Gaps 0;
QY	8	acggcgaccacgaagaacgcgattgctcagcagatgcattgcctgcgctggtgatact	67	
Db	410	AAAGGGACCGCAAAACATTAATTCACGAAGCATATCATCATCATCATTCATTCAGCTAGGGGGTCTACT	469	









OK NCBI\_TaxID=9823;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RA Kijas J.M.H., Andersson L.;  
 RT "A phylogenetic study of the origin of the domestic pig estimated from  
 the near complete mtDNA genome."  
 RL Submitted (SEP-2000) to the EMBL/GenBank/DBJ databases.  
 DR EMBL; AF304201; AAG28196.1; -  
 DR InterPro; IPR000298; CytC\_oxdse\_III.  
 DR Pfam; PF00510; COX3; 1.  
 DR PROSITE; PS50253; COX3; 1.  
 KW Mitochondrion.  
 SQ SEQUENCE 267 AA; 30462 MW; 95C558C54F6C0E3F CRC64;

alignment\_scores:  
 Quality: 84.00 Length: 37  
 Ratio: 3.231 Gaps: 0  
 Percent Similarity: 70.270 Percent Identity: 51.351

alignment\_block:  
 US-09-712-768-5 x Q9G7T0 ..

Align seg 1/1 to: Q9G7T0 from: 1 to: 267

```

1 ATGCTCCAGCGGCGACGCAAGAAACCGGATTCGCTAGCGATTGCCAT 50
   ::::|||||
151 LeuMetGluGlyAspArgLysHisMetIleGlnAlaLeuSerIleThrII 167
   ::::|||||
51 CGGCGCTTGCGTGCATCTTACCTGTGCCAAGCCTATGAATATTGAAA 100
   ::::|||||
167 eAlaLeuGlyValTyrPheThrLeuLeuGlnAlaSerGluTyrTyrGluA 184
   ::::|||||
101 TCGTCCATACC 111
   |||
184 laProPheThr 187

```

seq\_name: sp\_organelle:Q9G7S0

seq\_documentation\_block:  
 ID Q9G7S0 PRELIMINARY; PRT; 267 AA.  
 AC Q9G7S0;  
 DT 01-MAR-2001 (TREMBlrel. 16, Created)  
 DT 01-MAR-2001 (TREMBlrel. 16, Last sequence update)  
 DT 01-JUN-2001 (TREMBlrel. 17, Last annotation update)  
 DE COIII.  
 GN COIII.  
 OS Sus scrofa (Pig).  
 OC Mitochondrion.  
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 OC Mammalia; Eutheria; Cetartiodactyla; Suina; Suidae; Sus.  
 OX NCBI\_TaxID=9823;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RA Kijas J.M.H., Andersson L.;  
 RT "A phylogenetic study of the origin of the domestic pig estimated from  
 the near complete mtDNA genome."  
 RL Submitted (SEP-2000) to the EMBL/GenBank/DBJ databases.  
 DR EMBL; AF304203; AAG28222.1; -  
 DR InterPro; IPR000298; CytC\_oxdse\_III.  
 DR Pfam; PF00510; COX3; 1.  
 DR PROSITE; PS50253; COX3; 1.  
 KW Mitochondrion.  
 SQ SEQUENCE 267 AA; 30454 MW; 471ABD9C294867E9 CRC64;

alignment\_scores:  
 Quality: 84.00 Length: 37  
 Ratio: 3.231 Gaps: 0  
 Percent Similarity: 70.270 Percent Identity: 51.351

alignment\_block:  
 US-09-712-768-5 x Q9G7S0 ..

Align seg 1/1 to: Q9G7S0 from: 1 to: 267

```

1 ATGCTCCAGCGGCGACGCAAGAAACCGGATTCGCTAGCGATTGCCAT 50
   ::::|||||
151 LeuMetGluGlyAspArgLysHisMetIleGlnAlaLeuSerIleThrII 167
   ::::|||||
51 CGGCGCTTGCGTGCATCTTACCTGTGCCAAGCCTATGAATATTGAAA 100
   ::::|||||
167 eAlaLeuGlyValTyrPheThrLeuLeuGlnAlaSerGluTyrTyrGluA 184
   ::::|||||
101 TCGTCCATACC 111
   |||
184 laProPheThr 187

```

seq\_name: sp\_organelle:Q9G290

seq\_documentation\_block:  
 ID Q9G290 PRELIMINARY; PRT; 278 AA.  
 AC Q9G290;

DT 01-MAR-2001 (TREMBlrel. 16, Created)  
 DT 01-MAR-2001 (TREMBlrel. 16, Last sequence update)  
 DT 01-JUN-2001 (TREMBlrel. 17, Last annotation update)  
 DE COIII.  
 GN COIII.  
 OS Sus scrofa (Pig).  
 OC Mitochondrion.  
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 OC Mammalia; Eutheria; Cetartiodactyla; Suina; Suidae; Sus.  
 OX NCBI\_TaxID=9823;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RA Kijas J.M.H., Andersson L.;  
 RT "A phylogenetic study of the origin of the domestic pig estimated from  
 the near complete mtDNA genome."  
 RL Submitted (SEP-2000) to the EMBL/GenBank/DBJ databases.  
 DR EMBL; AF304202; AAG28209.1; -  
 DR EMBL; AF304200; AAG28183.1; -  
 DR InterPro; IPR000298; CytC\_oxdse\_III.  
 DR Pfam; PF00510; COX3; 1.  
 DR PROSITE; PS50253; COX3; 1.  
 KW Mitochondrion.  
 SQ SEQUENCE 278 AA; 31927 MW; 25A60814452F5A2B CRC64;

alignment\_scores:  
 Quality: 84.00 Length: 37  
 Ratio: 3.231 Gaps: 0  
 Percent Similarity: 70.270 Percent Identity: 51.351

alignment\_block:  
US-09-712-768-5 x Q9G290 ..

Align seg 1/1 to: Q9G290 from: 1 to: 278

```

1 ATGCTCCAGCGGCGACGCAAGAAACCGGATTCGCTAGCGATTGCCAT 50
   ::::|||||
151 LeuMetGluGlyAspArgLysHisMetIleGlnAlaLeuSerIleThrII 167
   ::::|||||
51 CGGCGCTTGCGTGCATCTTACCTGTGCCAAGCCTATGAATATTGAAA 100
   ::::|||||
167 eAlaLeuGlyValTyrPheThrLeuLeuGlnAlaSerGluTyrTyrGluA 184
   ::::|||||
101 TCGTCCATACC 111
   |||
184 laProPheThr 187

```

seq\_name: sp\_organelle:Q9T244

seq\_documentation\_block:  
 ID Q9T244 PRELIMINARY; PRT; 305 AA.  
 AC Q9T244;

DT 01-MAY-2000 (TREMBlrel. 13, Created)

```

alignment_scores:      Length: 33
                        Quality: 84.00
                        Ratio: 3.111
                        Percent Similarity: 81.818      Percent Identity: 54.545

alignment_block:
  05-09-712-768-5 x Q97244  ..

Align seg 1/1 to: Q97244 from: 1 to: 305

1  ATCTGTCACGGCGACCCGAGAAACCGGAGCTTGGCTTGGCATTTGCCAT 50
   |||||  ||||| ||||| ||||| ||||| ||||| ||||| |||||
194  lilevalphgglYAsPArGlyAsnAlAlIleIleSerIleIleIleIleIle 210
   ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
51  CGGCGCTTGGCTGGATCTTACCCTGTGCCAAGCGTATGANTTATTCGA 99
   |  ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
210  eleulnAlAlAphephepSerIleIleGlnAlATyrlGlyIleGlu 226

seq_name: sp-organella:Q9B629

seq_documentation_block:
ID_09B629      PRELIMINARY;      PRT;      261 AA.
AC  Q9B629:
DT  01-JUN-2001 (TRENBIrel. 17, Created)
DT  01-JUN-2001 (TRENBIrel. 17, Last sequence update)
DT  01-JUN-2001 (TRENBIrel. 17, Last annotation update)
DE  CYTOCHROME C OXIDASE SUBUNIT III.
GN  COXIII.
OS  Anguilla japonica (Japanese eel).
OC  Mitochondrion.
OC  Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC  Actinopterygii; Neopterygii; Teleostei; Anguilliformes; Anguillidae;

```

```

OX Anguillidae; Anguilla.
RX NCBL_TaxID=7937;
RN [1]
RP SEQUENCE FROM N.A., Nishida M.;
RA Inoue J.G., Miya M., submitted (FEB-2000) to the EMBL/GenBank/DBJ databases.
RL Submitted (FEB-2000) to the EMBL/GenBank/DBJ databases.
RN [2]
RP SEQUENCE FROM N.A.,
RA Inoue J.G., Miya M., Aoyama J., Ishikawa S., Tsukamoto K., Nishida M.;
RT "Complete Mitochondrial DNA Sequence of the Japanese Fel, Anguilla
   japonica."
RI Fisheries Sci. 67:118-125(2001).
RD Fisheries Sci. 67:118-125(2001).
DR EMBL; AB038556; BAB39718.1; -.
KV Mitochondrion.
SQ SEQUENCE 261 AA; 29717 MW; D6317E3C76DA48A7 CRC64;

alignment_scores:
    Quality:      83.00          Length:       37
              Ratio:     2.964           Gaps:         0
Percent Similarity: 75.676        Percent Identity: 48.649

alignment_block:
US-09-712-768-5 x Q9B629 ..

Align seg 1/1 to: Q9B629 from: 1 to: 261

1 ATGCTCCACGGCGACGCAGAAAACC GGATTGCCCTTACGATTTGCCAT 50
11 |||||||:|||||||::|::|:|||||::|:|||||:|||||::|:
151 lilemetgiuglunargylsginalailglnserleutrinrl 167
51 CGGCGTCGTGCGATCTTTACCCCTGTGCCCACTCATGTAATAATGAGA 100
1 | |||||::| ||||| ||||| ||||| ||||| ||||| |||||
167 elleluenlypheyrphenrleuleugnalmemetgiuryrygua 184
101 TCGTCCATACC 111
184 laprobether 187

seq_name: sp_organelle:Q9GAU2

seq_documentation_block:
ID Q9GAU2 PRELIMINARY; PRT: 260 AA.

AC Q9GAU2;
DT 01-MAR-2001 (TREMBLE), 16, Created
DT 01-MAR-2001 (TREMBLE), 16, Last sequence update)
DT 01-JUN-2001 (TREMBLE), 17, Last annotation update)
DE CYTOCHROME C OXIDASE POLYPEPTIDE II.
OS Rana sylvatica (wood frog).
OG Mitochondrion.
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Amphibia; Batrachia; Anura; Neobatrachia; Ranioidea; Rana.
OX NCBL_TaxID=45438;
RN [1]
RP SEQUENCE FROM N.A.
RC TISSUE=BRAIN;
RA Wu S.-B., Storey K.B.;
RT "Up-regulation of mitochondrial ATPase genes under freezing exposure
   and anaoxic treatment by wood frog (Rana sylvatica).";
RL Submitted (Aug-1999) to the EMBL/GenBank/DBJ databases.
DR EMBL; AF115976; AGS33066.1; -.
DR InterPro; IPR000298; Cytochrome_c_oxdase_fam; PF00510; COX3; 1.
DR PROSITE; PS50253; COX3; 1.
KV Mitochondrion.
SQ SEQUENCE 260 AA; 29565 MW; 5639DC103F2CD186 CRC64;

alignment_scores:
    quality:      82.00          length:       37
              ratio:    3.154           gaps:         0
percent similarity: 70.270        percent identity: 51.351

```

## alignment\_block:

US-09-712-768-5 x Q9GAU2

Align seg 1/1 to: Q9GAU2 from: 1 to: 260

```

1 ATGCTCCAGCGGCGGCAAGAAACCGGATGCGCTAGCGATGGCCAT 50
|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
151 llemetcinalaspargylsglythrleuglnalaleuthrleuthr 167
51 CGGCTTGCTGGATCTTACCTGTGCGCAAGCCCTATGATATATGAAA 100
|||||:|||||:|||||:|||||:|||||:|||||:|||||:
167 ethleuglyleutyrrpethrleuenglinalamecglutyrtyglua 184
101 TCGTCATACC 111
|||
184 laPropher 187

```

seq\_name: sp\_organelle:Q9B614

## seq\_documentation\_block:

```

ID Q9B614 PRELIMINARY: PRT: 261 AA.
AC Q9B614;
DT 01-JUN-2001 (TREMBlrel. 17, Created)
DT 01-JUN-2001 (TREMBlrel. 17, Last sequence update)
DE CYTOCHROME C OXIDASE SUBUNIT III.
GN CO III.
OS Carassius auratus (Goldfish).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Actinopterygii; Neopterygii; Teleostei; Euteleostei; Ostariophysi;
OX NCBI_TaxID=7957;
RN [1]
RP SEQUENCE FROM N.A.
RA Murakami M.;
RT "Carassius auratus cv.let1 mitochondrial DNA, complete sequence.";
DR EMBL; AB045144; BAB40354.1; -
KW Mitochondrion.
SQ SEQUENCE 261 AA; 29658 MW; D0DD83AD5BE184FF CRC64;

```

## alignment\_scores:

Quality:	Ratio:	Length:
82.00	3.037	37
Percent Similarity: 72.973	Percent Identity: 48.649	Gaps: 0

## alignment\_block:

US-09-712-768-5 x Q9B614

Align seg 1/1 to: Q9B614 from: 1 to: 261

```

1 ATGCTCCAGCGGCGGCAAGAAACCGGATGCGCTAGCGATGGCCAT 50
|||||:|||||:|||||:|||||:|||||:|||||:|||||:
151 llemetcinalaspargylsglnalaleuinserteuAlaleuthr 167
51 CGGCTTGCTGGATCTTACCTGTGCGCAAGCCCTATGATATATGAAA 100
|||||:|||||:|||||:|||||:|||||:|||||:|||||:
167 ethleuglyleutyrrpethrleuenglinalamecglutyrtyglua 184
101 TCGTCATACC 111
|||
184 laPropher 187

```

seq\_name: sp\_organelle:Q9ZY39

## seq\_documentation\_block:

```

ID Q9ZY39 PRELIMINARY: PRT: 261 AA.
AC Q9ZY39;
DT 01-MAY-1999 (TREMBlrel. 10, Created)
DT 01-MAY-1999 (TREMBlrel. 10, Last sequence update)
DT 01-JUN-2001 (TREMBlrel. 17, Last annotation update)

```

```

DE CYTOCHROME C OXIDASE SUBUNIT III.
GN COX-III.
OS Raja radiata.
OG Mitochondrion.
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Chondrichthyes;
OC Elasmobranchii; Squalae; Hypnosqualae; Pristiogaster; Batoidae;
OC Rajiformes; Rajidae; Raja.
OX NCBI_TaxID=27589;
RN [1]
RP SEQUENCE FROM N.A.
RA Rasmussen A.-S.;
RC MEDLINE=99162577; PubMed=10051614;
RX "Molecular studies suggest that cartilaginous fishes have a terminal
RT position in the piscine tree."
RL Proc. Natl. Acad. Sci. U.S.A. 96:2177-2182(1999).
RN [2]
RP SEQUENCE FROM N.A.
RA Rasmussen A.-S.;
RC TISSUE=LIVER;
RX Submitted (NOV-1998) to the EMBL/GenBank/DBJ databases.
DR EMBL; AF106038; AAD19935.1; -.
DR HSSP; P18402; 1FFT.
DR InterPro; IPR000298; Cytochrome_c_oxidase_subunit_III.
DR Pfam; PF00510; COX3; 1.
DR PROSITE; PS50253; COX3; 1.
KW Mitochondrion.
SQ SEQUENCE 261 AA; 29600 MW; 6A9B1A6DC779F788 CRC64;

```

## alignment\_scores:

Quality:	Ratio:	Length:
81.00	3.000	37
Percent Similarity: 72.973	Percent Identity: 45.946	Gaps: 0

## alignment\_block:

US-09-712-768-5 x Q9ZY39

Align seg 1/1 to: Q9ZY39 from: 1 to: 261

```

1 ATGCTCCAGCGGCGGCAAGAAACCGGATGCGCTAGCGATGGCCAT 50
|||||:|||||:|||||:|||||:|||||:|||||:|||||:
151 llemetcinalaspargylsglnalaleuinserteuAlaleuthr 167
51 CGGCTTGCTGGATCTTACCTGTGCGCAAGCCCTATGATATATGAAA 100
|||||:|||||:|||||:|||||:|||||:|||||:|||||:
167 ethleuglyleutyrrpethrleuenglinalamecglutyrtyglua 184
101 TCGTCATACC 111
|||
184 laPropher 187

```

seq\_name: sp\_organelle:Q9XN34

## seq\_documentation\_block:

```

ID Q9XN34 PRELIMINARY: PRT: 261 AA.
AC Q9XN34;
DT 01-NOV-1999 (TREMBlrel. 12, Created)
DT 01-NOV-1999 (TREMBlrel. 12, Last sequence update)
DT 01-JUN-2001 (TREMBlrel. 17, Last annotation update)
DE CYTOCHROME OXIDASE SUBUNIT 3.
GN COIII.
OS Salvelinus fontinalis (Brook trout) (Brook char).
OG Mitochondrion.
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Actinopterygii; Neopterygii; Teleostei; Euteleostei;
OC Protacanthopterygii; Salmoniformes; Salmonidae; Salvelinus.
OX NCBI_TaxID=8038;
RN [1]
RP SEQUENCE FROM N.A.
RA Dolton S., Blier P.U., Bernatchez L.;
RT "A comparative analysis of complete sequence of mitochondrial genome
between brook char (Salvelinus fontinalis) and arctic char (S.

```

RT alpines).  
 RL Submitted (MAY-1999) to the EMBL/GenBank/DBJ databases.  
 DR EMBL: AF154850; AAD41378.1; .  
 DR HSBP: P18402; 1FFT.  
 DR InterPro: IPR000298; Cytc\_oxdse\_III.  
 DR Pfam: PF00510; COX3; 1.  
 DR PROSITE: PS50253; COX3; 1.  
 KW Mitochondrion.  
 SQ SEQUENCE 261 AA; 29721 MW; 9EAAAC788FE00316 CRC64;

alignment\_scores:  
 Quality: 81.00 Length: 37  
 Ratio: 2.793 Gaps: 0  
 Percent Similarity: 78.378 Percent Identity: 45.946

alignment\_block:  
 US-09-712-768-5 x Q9XN34 ..

Align seg 1/1 to: Q9XN34 from: 1 to: 261

```

1 ATGCTCACGGCGACCGCAAGAAACCGCATTTGGCTTACGATTCGCAT 50
|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
151 ILeMeTgLuGlYgLnrgYsGlnThrValGlnAlaLeuThrLeuThrII 167
51 CGGCTTGGCTGATCTTTACCCCTGTGCGCAAGCCTATGATATATGAAA 100
| |||||: |||||: |||||: |||||: |||||: |||||: |||||:
167 eLeuLeuGlyPheThrPheLeuGlnGlyMeTgLuTyTrYgLuA 184
101 TCGTCCATACC 111
|||
184 laProPheThr 187

```

seq\_name: sp\_organelle:Q9XN26

seq\_documentation\_block:  
 ID Q9XN26 PRELIMINARY; PRT; 261 AA.  
 AC Q9XN26;  
 DT 01-NOV-1999 (TREMblrel. 12, Created)  
 DT 01-NOV-1999 (TREMblrel. 12, Last sequence update)  
 DT 01-JUN-2001 (TREMblrel. 17, Last annotation update)  
 DE CYTOCHROME OXIDASE SUBUNIT 3.  
 GN COXII.  
 OS Salvelinus alpinus (Arctic char).  
 OG Mitochondrion.  
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 OC Actinopterygii; Neopterygii; Teleostei; Euteleostei;  
 OC Protacanthopterygii; Salmoniformes; Salmonidae; Salvelinus.  
 OC NCBITaxID=8036;  
 OX [1]  
 RN SEQUENCE FROM N.A.  
 RP Doliron S., Blier P.U., Bernatchez L.;  
 RA "A comparative analysis of complete sequence of mitochondrial genome  
 RT between brook char (salvelinus fontinalis) and arctic char (S.  
 RT alpines).";  
 RT Submitted (MAY-1999) to the EMBL/GenBank/DBJ databases.  
 RL EMBL: AF154851; AAD41391.1; .  
 DR HSBP: P18402; 1FFT.  
 DR InterPro: IPR000298; Cytc\_oxdse\_III.  
 DR Pfam: PF00510; COX3; 1.  
 DR PROSITE: PS50253; COX3; 1.  
 KW Mitochondrion.  
 SQ SEQUENCE 261 AA; 29751 MW; A96564CE40C15BA9 CRC64;

alignment\_scores:  
 Quality: 81.00 Length: 37  
 Ratio: 2.793 Gaps: 0  
 Percent Similarity: 78.378 Percent Identity: 45.946

alignment\_block:  
 US-09-712-768-5 x Q9XN26 ..

Align seg 1/1 to: Q9XN26 from: 1 to: 261

```

1 ATGCTCACGGCGACCGCAAGAAACCGCATTTGGCTTACGATTCGCAT 50
|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
151 ILeMeTgLuGlYgLnrgYsGlnThrValGlnAlaLeuThrLeuThrII 167
51 CGGCTTGGCTGATCTTTACCCCTGTGCGCAAGCCTATGATATATGAAA 100
| |||||: |||||: |||||: |||||: |||||: |||||: |||||:
167 eLeuLeuGlyPheThrPheLeuGlnGlyMeTgLuTyTrYgLuA 184
101 TCGTCCATACC 111
|||
184 laProPheThr 187

```

seq\_name: sp\_organelle:Q9G6Q7

seq\_documentation\_block:  
 ID Q9G6Q7 PRELIMINARY; PRT; 261 AA.  
 AC Q9G6Q7;  
 DT 01-MAR-2001 (TREMblrel. 16, Created)  
 DT 01-MAR-2001 (TREMblrel. 16, Last sequence update)  
 DT 01-JUN-2001 (TREMblrel. 17, Last annotation update)  
 DE CYTOCHROME C OXIDASE SUBUNIT III.  
 GN COXIII.  
 OS Diplophos taenia.  
 OG Mitochondrion.  
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 OC Actinopterygii; Neopterygii; Teleostei; Euteleostei; Neoteleostei;  
 OC Stomiliformes; Gonostomatidae; Diplophos.  
 OC NCBITaxID=91975;  
 OX [1]  
 RN SEQUENCE FROM N.A.  
 RP Miya M., Nishida M.;  
 RA "Setting a new stage for the teleostean molecular systematics: A  
 RT mtogenomic approach using long PCR and versatile primers.";  
 RT Submitted (NOV-1999) to the EMBL/GenBank/DBJ databases.  
 RL EMBL: AB034825; BAB20726.1; .  
 DR InterPro: IPR000298; Cytc\_oxdse\_III.  
 DR Pfam: PF00510; COX3; 1.  
 DR PROSITE: PS50253; COX3; 1.  
 KW Mitochondrion.  
 SQ SEQUENCE 261 AA; 29585 MW; 694856C933298B77 CRC64;

alignment\_scores:  
 Quality: 81.00 Length: 37  
 Ratio: 3.000 Gaps: 0  
 Percent Similarity: 72.973 Percent Identity: 48.649

alignment\_block:  
 US-09-712-768-5 x Q9G6Q7 ..

Align seg 1/1 to: Q9G6Q7 from: 1 to: 261

```

1 ATGCTCACGGCGACCGCAAGAAACCGCATTTGGCTTACGATTCGCAT 50
|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
151 ILeMeTgLuGlYgLnrgYsGlnAlaLeuThrLeuThrII 167
51 CGGCTTGGCTGATCTTTACCCCTGTGCGCAAGCCTATGATATATGAAA 100
| |||||: |||||: |||||: |||||: |||||: |||||: |||||:
167 eLeuLeuGlyPheThrPheLeuGlnGlyMeTgLuTyTrYgLuA 184
101 TCGTCCATACC 111
|||
184 laProPheThr 187

```

seq\_name: sp\_organelle:Q9G387

seq\_documentation\_block:  
 ID Q9G387 PRELIMINARY; PRT; 261 AA.  
 AC Q9G387;  
 DT 01-MAR-2001 (TREMblrel. 16, Created)  
 DT 01-MAR-2001 (TREMblrel. 16, Last sequence update)





101 TCGTCATACC 111  
111  
184 laPropherhr 187

---



Query Match	27.2%	Score 31; DB 2;	Length 1813;
Best Similarity	56.3%	Pred. No. 0.051;	
Best Local			
Matches 58; Conservative	0;	Mismatches 45;	Indels 0;
			Gaps 0



Thu Jan 10 08:15:42 2002

us-09-712-768-5.rni

ADDRESSER: Ciba-Geigy Corporation  
STREET: 7 Skyline Drive  
CITY: Hawthorne  
STATE: NY  
COUNTRY: USA  
ZIP: 10532  
COMPUTER READABLE FORM:  
MEDIUM TYPE: Floppy disk  
COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: Patentin Release #1.0, Version #1.25  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/08/258,261B  
FILING DATE: 08-JUN-1994  
CLASSIFICATION: 800  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: US 08/457,205  
FILING DATE: 01-JUN-1995  
ATTORNEY/AGENT INFORMATION:  
NAME: Elmer, James Scott  
REGISTRATION NUMBER: 36,129  
REFERENCE/DOCKET NUMBER: CGC 1506/CIP3  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: 919-541-8614  
TELEFAX: 919-541-8689  
INFORMATION FOR SEQ ID NO: 6:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 28958 base pairs  
TYPE: nucleic acid  
STRANDEDNESS: single  
TOPOLOGY: linear  
MOLECULE TYPE: DNA (genomic)  
HYPOTHETICAL: NO  
ANTI-SENSE: NO  
US-08-258-261B-6

Query Match 22.1%; Score 25.2; DB 1; Length 28958;  
Best Local Similarity 51.8%; Pred. No. 18;  
Matches 57; Conservative 0; Mismatches 53; Indels 0; Gaps 0;

OY 3 cgtccagcgacgcaagaacacgcatggtgctgagcgtgcatgagcgtgctgctg 62  
DB 12768 CATCGCCGACGCGCCGATGCTGCCCATGCCCGCTGCTTCGCCCAAGCCGC 12827

OY 63 gatcttaccctgtgccaagcctatgaatattatgaatcgtccatccg 112  
DB 12828 CAGCGTCCCATCATCTATCTACCGCCTACTATGAGCTGCGATCTCG 12877

RESULT 6  
US-08-456-837-6  
Sequence 6, Application US/08456837  
Patent No. 5643774  
GENERAL INFORMATION:  
APPLICANT: Schupp, Thomas  
APPLICANT: Ligon, James M.  
APPLICANT: Beck, James Joseph  
APPLICANT: Hill, Dwight Steven  
APPLICANT: Ryals, John Andrew  
APPLICANT: Gaffney, Thomas Deane  
APPLICANT: Lam, Stephen Ting  
APPLICANT: Hammer, Phillip E.  
APPLICANT: Ukner, Scott Joseph  
TITLE OF INVENTION: Genes for the synthesis of  
TITLE OF INVENTION: antipathogenic substances  
NUMBER OF SEQUENCES: 22  
CORRESPONDENCE ADDRESS:  
ADDRESSER: Ciba-Geigy Corporation  
STREET: 7 Skyline Drive  
CITY: Hawthorne  
STATE: NY  
COUNTRY: USA

ZIP: 10532  
COMPUTER READABLE FORM:  
MEDIUM TYPE: Floppy disk  
COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: Patentin Release #1.0, Version #1.25  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/08/456,837  
FILING DATE: 01-JUN-1995  
CLASSIFICATION: 435  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: 08/457,205  
FILING DATE: 01-JUN-1995  
APPLICATION NUMBER: 08/258,261  
FILING DATE: 08-JUN-1994  
ATTORNEY/AGENT INFORMATION:  
NAME: Elmer, James Scott  
REGISTRATION NUMBER: 36,129  
REFERENCE/DOCKET NUMBER: CGC 1506/CIP3  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: 919-541-8614  
TELEFAX: 919-541-8689  
INFORMATION FOR SEQ ID NO: 6:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 28958 base pairs  
TYPE: nucleic acid  
STRANDEDNESS: single  
TOPOLOGY: linear  
MOLECULE TYPE: DNA (genomic)  
HYPOTHETICAL: NO  
ANTI-SENSE: NO  
US-08-456-837-6

Query Match 22.1%; Score 25.2; DB 1; Length 28958;  
Best Local Similarity 51.8%; Pred. No. 18;  
Matches 57; Conservative 0; Mismatches 53; Indels 0; Gaps 0;

OY 3 cgtccagcgacgcaagaacacgcatggtgctgagcgtgcatgagcgtgctgctg 62  
DB 12768 CATCGCCGACGCGCCGATGCTGCCCATGCCCGCTGCTTCGCCCAAGCCGC 12827

OY 63 gatcttaccctgtgccaagcctatgaatattatgaatcgtccatccg 112  
DB 12828 CAGCGTCCCATCATCTATCTACCGCCTACTATGAGCTGCGATCTCG 12877

RESULT 7  
US-08-457-342-6  
Sequence 6, Application US/08457342  
Patent No. 5662898  
GENERAL INFORMATION:  
APPLICANT: Schupp, Thomas  
APPLICANT: Ligon, James M.  
APPLICANT: Beck, James Joseph  
APPLICANT: Hill, Dwight Steven  
APPLICANT: Ryals, John Andrew  
APPLICANT: Gaffney, Thomas Deane  
APPLICANT: Lam, Stephen Ting  
APPLICANT: Hammer, Phillip E.  
APPLICANT: Ukner, Scott Joseph  
TITLE OF INVENTION: Genes for the synthesis of  
TITLE OF INVENTION: antipathogenic substances  
NUMBER OF SEQUENCES: 22  
CORRESPONDENCE ADDRESS:  
ADDRESSER: Ciba-Geigy Corporation  
STREET: 7 Skyline Drive  
CITY: Hawthorne  
STATE: NY  
COUNTRY: USA  
COMPUTER READABLE FORM:  
MEDIUM TYPE: Floppy disk

COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: Patentin Release #1.0, Version #1.25  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/08/457,342  
FILING DATE: 01-JUN-1995  
CLASSIFICATION: 424  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: US 08/457,205  
FILING DATE: 01-JUN-1995  
APPLICATION NUMBER: 08/258,261  
FILING DATE: 08-JUN-1994  
ATTORNEY/AGENT INFORMATION:  
NAME: Elmer, James Scott  
REGISTRATION NUMBER: 36,129  
REFERENCE/DOCKET NUMBER: CGC 1506/CIP3  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: 919-541-8614  
TELEFAX: 919-541-8689  
INFORMATION FOR SEQ ID NO: 6:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 28958 base pairs  
TYPE: nucleic acid  
STRANDEDNESS: single  
TOPOLOGY: linear  
MOLECULE TYPE: DNA (genomic)  
HYPOTHETICAL: NO  
ANTI-SENSE: NO  
US-08-457-342-6

Query Match 22.1%; Score 25.2; DB 1; Length 28958;  
Best Local Similarity 51.8%; Pred. No. 18;  
Matches 57; Conservative 0; Mismatches 53; Indels 0; Gaps 0;  
QY 3 cgtcacggcgacccaagaacacgcatgagcagcattgacatcgcccttgctg 62  
DB 12768 CATCCCGACGCCCGCATGATCTGCCCATCCCGCCTGCTCTTCCCAAGCCGC 12827  
QY 63 gatcttaccctgtgcgaagcctatgaatatgaatcgatcattacgg 112  
DB 12828 CAGCGTCCCATCATCTATCTACCGCCCTACTATGACTGCTCATCTCG 12877

RESULT 8  
US-08-457-646A-6  
Sequence 6, Application US/08457646A  
Patent No. 5679560  
GENERAL INFORMATION:  
APPLICANT: Schupp, Thomas  
APPLICANT: Lyon, James M.  
APPLICANT: Beck, James Joseph  
APPLICANT: Hill, Dwight Steven  
APPLICANT: Ryals, John Andrew  
APPLICANT: Gaffney, Thomas Deane  
APPLICANT: Lam, Stephen Tling  
APPLICANT: Hammer, Phillip E.  
TITLE OF INVENTION: Genes for the synthesis of  
NUMBER OF INVENTION: antipathogenic substances  
NUMBER OF SEQUENCES: 22  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: Ciba-Geigy Corporation  
STREET: 7 Skyline Drive  
CITY: Hawthorne  
STATE: NY  
COUNTRY: USA  
ZIP: 10532  
COMPUTER READABLE FORM:  
MEDIUM TYPE: Floppy disk  
COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: Patentin Release #1.0, Version #1.25  
FILING DATE: 01-JUN-1995

CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/08/457,646A  
FILING DATE: 01-JUN-1995  
CLASSIFICATION: 530  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: US 08/457,205  
FILING DATE: 01-JUN-1995  
APPLICATION NUMBER: 08/258,261  
FILING DATE: 08-JUN-1994  
ATTORNEY/AGENT INFORMATION:  
NAME: Elmer, James Scott  
REGISTRATION NUMBER: 36,129  
REFERENCE/DOCKET NUMBER: CGC 1506/CIP3  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: 919-541-8614  
TELEFAX: 919-541-8689  
INFORMATION FOR SEQ ID NO: 6:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 28958 base pairs  
TYPE: nucleic acid  
STRANDEDNESS: single  
TOPOLOGY: linear  
MOLECULE TYPE: DNA (genomic)  
HYPOTHETICAL: NO  
ANTI-SENSE: NO  
US-08-457-646A-6

Query Match 22.1%; Score 25.2; DB 1; Length 28958;  
Best Local Similarity 51.8%; Pred. No. 18;  
Matches 57; Conservative 0; Mismatches 53; Indels 0; Gaps 0;  
QY 3 cgtcacggcgacccaagaacacgcatgagcagcattgacatcgcccttgctg 62  
DB 12768 CATCCCGACGCCCGCATGATCTGCCCATCCCGCCTGCTCTTCCCAAGCCGC 12827  
QY 63 gatcttaccctgtgcgaagcctatgaatatgaatcgatcattacgg 112  
DB 12828 CAGCGTCCCATCATCTATCTACCGCCCTACTATGACTGCTCATCTCG 12877

RESULT 9  
US-08-458-076A-6  
Sequence 6, Application US/08458076A  
Patent No. 5698425  
GENERAL INFORMATION:  
APPLICANT: Schupp, Thomas  
APPLICANT: Lyon, James M.  
APPLICANT: Beck, James Joseph  
APPLICANT: Hill, Dwight Steven  
APPLICANT: Ryals, John Andrew  
APPLICANT: Gaffney, Thomas Deane  
APPLICANT: Lam, Stephen Tling  
APPLICANT: Hammer, Phillip E.  
TITLE OF INVENTION: Genes for the synthesis of  
NUMBER OF INVENTION: antipathogenic substances  
NUMBER OF SEQUENCES: 22  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: Ciba-Geigy Corporation  
STREET: 7 Skyline Drive  
CITY: Hawthorne  
STATE: NY  
COUNTRY: USA  
ZIP: 10532  
COMPUTER READABLE FORM:  
MEDIUM TYPE: Floppy disk  
COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: Patentin Release #1.0, Version #1.25  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/08/458,076A  
FILING DATE: 01-JUN-1995

```

ATTORNEY/AGENT INFORMATION:
1: NAME: Meigs, J. Timothy
2: REGISTRATION NUMBER: 38,241
3: REFERENCE/DOCKET NUMBER: 1506/CIP6
4: TELECOMMUNICATION INFORMATION:
5: TELEPHONE: (919) 541-8587
6: TELEFAX: (919) 541-8689
7: INFORMATION FOR SEQ ID NO: 4:
8: SEQUENCE CHARACTERISTICS:
9: LENGTH: 28958 base pairs
10: TYPE: nucleic acid
11: STRANDEDNESS: single
12: TOPOLOGY: linear
13: MOLECULE TYPE: DNA (genomic)
14: ORIGINAL SOURCE:
15: ORGANISM: Sorangium cellulosum
16: IMMEDIATE SOURCE:
17: CLONE: p98/1
18:
19: US-08-764-233A-4
20:
21:
22:
23:
24:
25:
26:
27:
28:
29:
30:
31:
32:
33:
34:
35:
36:
37:
38:
39:
40:
41:
42:
43:
44:
45:
46:
47:
48:
49:
50:
51:
52:
53:
54:
55:
56:
57:
58:
59:
60:
61:
62:
63:
64:
65:
66:
67:
68:
69:
70:
71:
72:
73:
74:
75:
76:
77:
78:
79:
80:
81:
82:
83:
84:
85:
86:
87:
88:
89:
90:
91:
92:
93:
94:
95:
96:
97:
98:
99:
100:
101:
102:
103:
104:
105:
106:
107:
108:
109:
110:
111:
112:
113:
114:
115:
116:
117:
118:
119:
120:
121:
122:
123:
124:
125:
126:
127:
128:
129:
130:
131:
132:
133:
134:
135:
136:
137:
138:
139:
140:
141:
142:
143:
144:
145:
146:
147:
148:
149:
150:
151:
152:
153:
154:
155:
156:
157:
158:
159:
160:
161:
162:
163:
164:
165:
166:
167:
168:
169:
170:
171:
172:
173:
174:
175:
176:
177:
178:
179:
180:
181:
182:
183:
184:
185:
186:
187:
188:
189:
190:
191:
192:
193:
194:
195:
196:
197:
198:
199:
200:
201:
202:
203:
204:
205:
206:
207:
208:
209:
210:
211:
212:
213:
214:
215:
216:
217:
218:
219:
220:
221:
222:
223:
224:
225:
226:
227:
228:
229:
230:
231:
232:
233:
234:
235:
236:
237:
238:
239:
240:
241:
242:
243:
244:
245:
246:
247:
248:
249:
250:
251:
252:
253:
254:
255:
256:
257:
258:
259:
260:
261:
262:
263:
264:
265:
266:
267:
268:
269:
270:
271:
272:
273:
274:
275:
276:
277:
278:
279:
280:
281:
282:
283:
284:
285:
286:
287:
288:
289:
290:
291:
292:
293:
294:
295:
296:
297:
298:
299:
300:
301:
302:
303:
304:
305:
306:
307:
308:
309:
310:
311:
312:
313:
314:
315:
316:
317:
318:
319:
320:
321:
322:
323:
324:
325:
326:
327:
328:
329:
330:
331:
332:
333:
334:
335:
336:
337:
338:
339:
340:
341:
342:
343:
344:
345:
346:
347:
348:
349:
350:
351:
352:
353:
354:
355:
356:
357:
358:
359:
360:
361:
362:
363:
364:
365:
366:
367:
368:
369:
370:
371:
372:
373:
374:
375:
376:
377:
378:
379:
380:
381:
382:
383:
384:
385:
386:
387:
388:
389:
390:
391:
392:
393:
394:
395:
396:
397:
398:
399:
400:
401:
402:
403:
404:
405:
406:
407:
408:
409:
410:
411:
412:
413:
414:
415:
416:
417:
418:
419:
420:
421:
422:
423:
424:
425:
426:
427:
428:
429:
430:
431:
432:
433:
434:
435:
436:
437:
438:
439:
440:
441:
442:
443:
444:
445:
446:
447:
448:
449:
450:
451:
452:
453:
454:
455:
456:
457:
458:
459:
460:
461:
462:
463:
464:
465:
466:
467:
468:
469:
470:
471:
472:
473:
474:
475:
476:
477:
478:
479:
480:
481:
482:
483:
484:
485:
486:
487:
488:
489:
490:
491:
492:
493:
494:
495:
496:
497:
498:
499:
500:
501:
502:
503:
504:
505:
506:
507:
508:
509:
510:
511:
512:
513:
514:
515:
516:
517:
518:
519:
520:
521:
522:
523:
524:
525:
526:
527:
528:
529:
530:
531:
532:
533:
534:
535:
536:
537:
538:
539:
540:
541:
542:
543:
544:
545:
546:
547:
548:
549:
550:
551:
552:
553:
554:
555:
556:
557:
558:
559:
560:
561:
562:
563:
564:
565:
566:
567:
568:
569:
570:
571:
572:
573:
574:
575:
576:
577:
578:
579:
580:
581:
582:
583:
584:
585:
586:
587:
588:
589:
590:
591:
592:
593:
594:
595:
596:
597:
598:
599:
600:
601:
602:
603:
604:
605:
606:
607:
608:
609:
610:
611:
612:
613:
614:
615:
616:
617:
618:
619:
620:
621:
622:
623:
624:
625:
626:
627:
628:
629:
630:
631:
632:
633:
634:
635:
636:
637:
638:
639:
640:
641:
642:
643:
644:
645:
646:
647:
648:
649:
650:
651:
652:
653:
654:
655:
656:
657:
658:
659:
660:
661:
662:
663:
664:
665:
666:
667:
668:
669:
670:
671:
672:
673:
674:
675:
676:
677:
678:
679:
680:
681:
682:
683:
684:
685:
686:
687:
688:
689:
690:
691:
692:
693:
694:
695:
696:
697:
698:
699:
700:
701:
702:
703:
704:
705:
706:
707:
708:
709:
710:
711:
712:
713:
714:
715:
716:
717:
718:
719:
720:
721:
722:
723:
724:
725:
726:
727:
728:
729:
730:
731:
732:
733:
734:
735:
736:
737:
738:
739:
740:
741:
742:
743:
744:
745:
746:
747:
748:
749:
750:
751:
752:
753:
754:
755:
756:
757:
758:
759:
760:
761:
762:
763:
764:
765:
766:
767:
768:
769:
770:
771:
772:
773:
774:
775:
776:
777:
778:
779:
780:
781:
782:
783:
784:
785:
786:
787:
788:
789:
790:
791:
792:
793:
794:
795:
796:
797:
798:
799:
800:
801:
802:
803:
804:
805:
806:

```

NAME: Elmer, James Scott  
 REGISTRATION NUMBER: 36,129  
 REFERENCE/DOCKET NUMBER: CGC 1506/CIP3  
 TELECOMMUNICATION INFORMATION:  
 TELEPHONE: 919-541-8614  
 INFORMATION FOR SEQ ID NO: 6:  
 SEQUENCE CHARACTERISTICS:  
 LENGTH: 28958 base pairs  
 TYPE: nucleic acid  
 STRANDEDNESS: single  
 TOPOLOGY: linear  
 MOLECULE TYPE: DNA (genomic)  
 HYPOTHETICAL: NO  
 ANTI-SENSE: NO  
 US-08-457-335A-6

Query Match  
 Best Local Similarity 22.1%; Score 25.2; DB 1; Length 28958;  
 Matches 57; Conservative 0; Mismatches 53; Indels 0; Gaps 0;

Db 12768 CATGCCGACGCCGCGATGTCGCCCATGCCCGACGCGTGTCTGCGCCCAAGCCG 12827

QY 63 gattcttaccctgtgccaaagcctatgaattatgaatgcgcataccg 112

Db 12828 CAGCGTCCCATCATCTATCTACCGCCTACTATGAGACTGTCGATCTCG 12877

RESULT 12  
 US-08-729-214-6

; Sequence 6, Application US/08729214  
 ; Patent No. 5817502

GENERAL INFORMATION:

APPLICANT: Ligon, James M.

APPLICANT: Hill, Dwight Steven

APPLICANT: Ryals, John Andrew

APPLICANT: Hammer, Phillip E.

APPLICANT: Van Pee, Karl-Heinz

APPLICANT: Kliner, Sabine

TITLE OF INVENTION: Genes for the synthesis of

NUMBER OF SEQUENCES: 27

CORRESPONDENCE ADDRESS:

ADDRESSEE: Ciba-Geigy Corporation

STREET: 520 White Plains Road

CITY: Tarrytown

STATE: NY

COUNTRY: USA

ZIP: 10591

COMPUTER READABLE FORM:

MEDIUM TYPE: Floppy disk

COMPUTER: IBM PC compatible

OPERATING SYSTEM: PC-DOS/MS-DOS

SOFTWARE: Patent Release #1.0, Version #1.25

CURRENT APPLICATION DATA:

APPLICATION NUMBER: US/08/729,214

FILING DATE: TBA

CLASSIFICATION: 435

ATTORNEY/AGENT INFORMATION:

NAME: Meigs, J. Timothy

REGISTRATION NUMBER: 38,241

REFERENCE/DOCKET NUMBER: CGC 1506/CIP5

TELECOMMUNICATION INFORMATION:

TELEPHONE: 919-541-8689

INFORMATION FOR SEQ ID NO: 6:

SEQUENCE CHARACTERISTICS:

LENGTH: 28958 base pairs

TYPE: nucleic acid

STRANDEDNESS: single

TOPOLOGY: linear  
 MOLECULE TYPE: DNA (genomic)  
 HYPOTHETICAL: NO  
 ANTI-SENSE: NO  
 US-08-729-214-6

Query Match  
 Best Local Similarity 22.1%; Score 25.2; DB 1; Length 28958;  
 Matches 57; Conservative 0; Mismatches 53; Indels 0; Gaps 0;

Db 12768 CATGCCGACGCCGCGATGTCGCCCATGCCCGACGCGTGTCTGCGCCCAAGCCG 12827

QY 63 gattcttaccctgtgccaaagcctatgaattatgaatgcgcataccg 112

Db 12828 CAGCGTCCCATCATCTATCTACCGCCTACTATGAGACTGTCGATCTCG 12877

RESULT 13

US-09-028-934-6  
 ; Sequence 6, Application US/09028934  
 ; Patent No. 6117670

GENERAL INFORMATION:

APPLICANT: Ligon, James M.

APPLICANT: Hill, Dwight S.

APPLICANT: Lam, Steven T.

APPLICANT: Hammer, Phillip E.

APPLICANT: Van Pee, Karl-Heinz

APPLICANT: Kliner, Sabine

TITLE OF INVENTION: Pyrolytic Biosynthesis Genes and Uses

NUMBER OF SEQUENCES: 37

CORRESPONDENCE ADDRESS:

ADDRESSEE: No. 6117670artis Corporation

STREET: 3054 Cornwallis Road

CITY: Research Triangle Park

STATE: NC

COUNTRY: USA

ZIP: 27709

COMPUTER READABLE FORM:

MEDIUM TYPE: Floppy disk

COMPUTER: IBM PC compatible

OPERATING SYSTEM: PC-DOS/MS-DOS

SOFTWARE: Patent Release #1.0, Version #1.30

CURRENT APPLICATION DATA:

APPLICATION NUMBER: US/09/028,934

FILING DATE: US/09/028,934

CLASSIFICATION:

PRIOR APPLICATION DATA:

APPLICATION NUMBER: US 08/729,214

FILING DATE: 09-OCT-1996

PRIOR APPLICATION DATA:

APPLICATION NUMBER: US 08/258,261

FILING DATE: 08-JUN-1994

ATTORNEY/AGENT INFORMATION:

NAME: Meigs, J. Timothy

REGISTRATION NUMBER: 38,241

REFERENCE/DOCKET NUMBER: CGC1506/CIP7

TELECOMMUNICATION INFORMATION:

TELEPHONE: 919-541-8689

INFORMATION FOR SEQ ID NO: 6:

SEQUENCE CHARACTERISTICS:

LENGTH: 28958 base pairs

TYPE: nucleic acid

STRANDEDNESS: single

TOPOLOGY: linear

MOLECULE TYPE: DNA (genomic)

HYPOTHETICAL: NO

ANTI-SENSE: NO



US-09-028-934-6

Query Match	22.1%;	Score 25.2;	DB 3;	Length 28958;
Best Local Similarity	51.8%;	Pred. No. 18;		
Matches 57; Conservative	0;	Mismatches 53;	Indels 0;	

QY 3 cgtccacgagcgaccgaagaaacgcgcatcttgcttagcattgccatcgaccttggctg 62  
| | | | | | | | | | | | | | | |  
Db 12768 CATGCCCGAGCCGCCGATGATCTGGCCCCCATCCCACGCGCTGTCTTCGCCCCAAGCCGC 12827

63 gatctttacccctgccaagccatgatatattatgaataatcgctccataccg 112  
 | | | | | | | | | | | | | | | | | | | | | | | |  
 Db 12828 CACGGTCCCCATCATCTATCTACACCGCCTACTAGGACTCGTGACTCTCG 12877

```

RESULT 14
US-08-764-233A-1
Sequence 1, Application US/08/64233A
Patent No. 5716849
GENERAL INFORMATION:
APPLICANT: Ligon, James M.
APPLICANT: Schupp, Thomas
APPLICANT: Beck, James J.
APPLICANT: Hill, Dwight S.
APPLICANT: Neff, Snezana
APPLICANT: Ryals, John A.
TITLE OF INVENTION: Genes For The Biosynthesis Of Soraphen
NUMBER OF SEQUENCES: 10
CORRESPONDENCE ADDRESS:
ADDRESSEE: Ciba-Geigy Corporation
STREET: 520 White Plains Road, P.O. Box 2005
CITY: Tarrytown
STATE: NY
COUNTRY: USA
ZIP: 10591
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: PatentIn Release #1.0, Version #1.30
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/764,233A
FILING DATE:
CLASSIFICATION: 435
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 08/729,214
FILING DATE: 09-OCT-1996
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 08/258,261
FILING DATE: 08-JUN-1994
ATTORNEY/AGENT INFORMATION:
NAME: Meigs, J. Timothy
REGISTRATION NUMBER: 38,241
REFERENCE/DOCKET NUMBER: 1506/CIP6
TELECOMMUNICATION INFORMATION:
TELEPHONE: (919) 541-8587
TELEFAX: (919) 541-8689
INFORMATION FOR SEQ ID NO: 1:
SEQUENCE CHARACTERISTICS:
LENGTH: 49377 base pairs
TYPE: nucleic acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: DNA (genomic)
ORGANISM: Sorangium cellulosum
IMMEDIATE SOURCE:
CLONE: p98/1, pJL3, and pVKM15
FEATURE:
NAME/KEY: misc.feature
LOCATION: 383..760
OTHER INFORMATION: /product= "Sorap"

```

1 OTHER INFORMATION: /note= "This gene encodes a protein that is highly homologous  
2 OTHER INFORMATION: /note= "The reductase domains of type I PKRS such as eryA from  
3 OTHER INFORMATION: Saccharopolyspora erythraea."  
4 FEATURE:  
5 NAME/KEY: misc\_feature  
6 LOCATION: 927..19874  
7 OTHER INFORMATION: /product= "Sora"  
8 OTHER INFORMATION: /note= "Gene product is highly homologous to type I PKRS t  
9 OTHER INFORMATION: are known to be involved in the synthesis of polyketide  
10 FEATURE:  
11 NAME/KEY: misc\_feature  
12 LOCATION: 942..7115  
13 OTHER INFORMATION: /product= "Module 1 of SORA"  
14 FEATURE:  
15 NAME/KEY: misc\_feature  
16 LOCATION: 7203..12884  
17 OTHER INFORMATION: /product= "Module 2 of SORA"  
18 FEATURE:  
19 NAME/KEY: misc\_feature  
20 LOCATION: 13455..19616  
21 OTHER INFORMATION: /product= "Module 3 of SORA"  
22 FEATURE:  
23 NAME/KEY: misc\_feature  
24 LOCATION: 19871..46318  
25 OTHER INFORMATION: /product= "SorB"  
26 OTHER INFORMATION: /note= "Gene product is highly homologous to type I PKRS ge  
27 FEATURE:  
28 NAME/KEY: misc\_feature  
29 LOCATION: 19870..24596  
30 OTHER INFORMATION: /product= "Module 1 of SorB"  
31 FEATURE:  
32 NAME/KEY: misc\_feature  
33 LOCATION: 24658..30820  
34 OTHER INFORMATION: /product= "Module 2 of SorB"  
35 FEATURE:  
36 NAME/KEY: misc\_feature  
37 LOCATION: 30881..35446  
38 OTHER INFORMATION: /product= "Module 3 of SorB"  
39 FEATURE:  
40 NAME/KEY: misc\_feature  
41 LOCATION: 35528..40114  
42 OTHER INFORMATION: /product= "Module 4 of SorB"  
43 FEATURE:  
44 NAME/KEY: misc\_feature  
45 LOCATION: 40190..46318  
46 OTHER INFORMATION: /product= "Module 5 of SorB"  
47 FEATURE:  
48 NAME/KEY: misc\_feature  
49 LOCATION: 46851..47891  
50 OTHER INFORMATION: /product= "SorM"  
51 OTHER INFORMATION: /note= "The protein encoded by the sorM gene is highly  
52 OTHER INFORMATION: homologous to the methyltransferase from Streptomyces  
53 OTHER INFORMATION: hygroscopicus that is involved in the synthesis of the  
54 OTHER INFORMATION: polyketide rapamycin."

US-08-764-233A-1

Query Match	22.1%	Score 25.2	DB 1	Length 43377
Best Local Similarity	51.8%	Pred. No. 22		
Matches 57; Conservative	0	Mismatches 53	Indels 0	Gaps 0

QY 3 cgtccacgagcgaaccgaanaaacgcgatctgctatgcattgcacatcgacttgctg 62  
| | | | | | | | | | | | | | | |  
Db 29071 CATGCCGAGCCCGCATGTATGCCCATCCCCACGCTGGTCTTCGCCCAAGCCC 29130

QY 63 gatctttacccctgccaagcctatgaatatattgaataatcgltcataccg 112  
| | | | | | | | | | | | | | | | | |  
Db 29131 CAGCGTCCCATCATCTATCTCAACGCCCTACTAATGGAATCGTCGATCTCG 29180

RESULT 15  
US-08-920-236-1/c

; Sequence 1, Application US/08920236  
; Patent No. 6225083  
; GENERAL INFORMATION:  
; APPLICANT: Hodgson, John E.  
; TITLE OF INVENTION: NOVEL FLSL  
; NUMBER OF SEQUENCES: 2  
; CORRESPONDENCE ADDRESS:  
; ADDRESSEE: Dechert Price & Rhoads  
; STREET: 4000 Bell Atlantic Tower, 1717 Arch Stre  
; CITY: Philadelphia  
; STATE: PA  
; COUNTRY: US  
; ZIP: 19103  
; COMPUTER READABLE FORM:  
; MEDIUM TYPE: Diskette  
; COMPUTER: IBM Compatible  
; OPERATING SYSTEM: DOS  
; SOFTWARE: FASTEO for Windows Version 2.0  
; CURRENT APPLICATION DATA:  
; APPLICATION NUMBER: US/08/920,236  
; FILING DATE:  
; CLASSIFICATION: 536  
; PRIOR APPLICATION DATA:  
; APPLICATION NUMBER:  
; FILING DATE:  
; ATTORNEY/AGENT INFORMATION:  
; NAME: Dickinson, Todd Q  
; REGISTRATION NUMBER: 28,354  
; REFERENCE/DOCKET NUMBER: P50556-1  
; TELECOMMUNICATION INFORMATION:  
; TELEPHONE: 215-994-2252  
; TELEFAX: 215-994-2222  
; TELEX:  
; INFORMATION FOR SEQ ID NO: 1:  
; SEQUENCE CHARACTERISTICS:  
; LENGTH: 315 base pairs  
; TYPE: nucleic acid  
; STRANDEDNESS: double  
; TOPOLOGY: linear  
; US-08-920-236-1

Query Match 21.9%; Score 25; DB 4; Length 315;  
Best Local Similarity 52.4%; Pred. No. 3.9;  
Matches 55; Conservative 0; Mismatches 50; Indels 0; Gaps 0;  
OY 2 tcgtccacgagcgacgaagaacccgagttgagcagatgcacatgcagccttgagt 61  
DB 260 TCTTCAACGCTTCGACAGTAATAGTTCATGTGACTCTTGTGGCATCGTCCAAATTGC 201  
OY 62 ggaatcttaccctgtgccaagcctatgaatatatgaatcgtcc 106  
DB 200 GTTTTCTTTCTCTATCTGCGCATGATTTTGTCAATCATTC 156

Search completed: January 10, 2002, 01:00:12  
Job time: 3180 sec

GenCore version 4.5  
Copyright (c) 1993 - 2000 CompuGen Ltd

OM nucleic - nucleic search, using sw model

Run on: January 10, 2002, 00:06:27 ; Search time 129.27 Seconds  
(without alignments)  
756.054 Million cell updates/sec

Title: US-09-712-768-5

Sequence: 1 atcgtccacgcgcaccgcaa.....atgaatcgtccataccgaa 114

Scoring code:	IDENTITY-NO
Gapop 10.0 ,	Gapext 1.0

Searched: 930621 seqs, 428662619 residues

Total number of hits satisfying chosen parameters: 1861242

```
Minimum DB seq length: 0
Maximum DB seq length: 20000000000
```

Post-processing: Minimum Match 0%  
Maximum Match 100%  
Listing first 45 summaries

1	./SID22/gcgcata./geneseq./geneseqn/NA1980.DAT.*
2	./SID22/gcgcata./geneseq./geneseqn/NA1981.DAT.*
3	./SID22/gcgcata./geneseq./geneseqn/NA1982.DAT.*
4	./SID22/gcgcata./geneseq./geneseqn/NA1983.DAT.*
5	./SID22/gcgcata./geneseq./geneseqn/NA1984.DAT.*
6	./SID22/gcgcata./geneseq./geneseqn/NA1985.DAT.*
7	./SID22/gcgcata./geneseq./geneseqn/NA1986.DAT.*
8	./SID22/gcgcata./geneseq./geneseqn/NA1987.DAT.*
9	./SID22/gcgcata./geneseq./geneseqn/NA1988.DAT.*
10	./SID22/gcgcata./geneseq./geneseqn/NA1989.DAT.*
11	./SID22/gcgcata./geneseq./geneseqn/NA1990.DAT.*
12	./SID22/gcgcata./geneseq./geneseqn/NA1991.DAT.*
13	./SID22/gcgcata./geneseq./geneseqn/NA1992.DAT.*
14	./SID22/gcgcata./geneseq./geneseqn/NA1993.DAT.*
15	./SID22/gcgcata./geneseq./geneseqn/NA1994.DAT.*
16	./SID22/gcgcata./geneseq./geneseqn/NA1995.DAT.*
17	./SID22/gcgcata./geneseq./geneseqn/NA1996.DAT.*
18	./SID22/gcgcata./geneseq./geneseqn/NA1997.DAT.*
19	./SID22/gcgcata./geneseq./geneseqn/NA1998.DAT.*
20	./SID22/gcgcata./geneseq./geneseqn/NA1999.DAT.*
21	./SID22/gcgcata./geneseq./geneseqn/NA2000.DAT.*
22	./SID22/gcgcata./geneseq./geneseqn/NA2001.DAT.*

**Pred. No.** is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed and is derived by analysis of the total score distribution.

## SUMMARIES

Result No.	Score	Query Match	Length	DB	ID	Description
1	114	100.0	114	22	AAA91491	G. oxydans cytoch
2	33.8	29.6	2580	21	AAD00811	Human membrane tr
3	31	27.2	1813	15	AAQ70903	Cold acclimatizat
4	27.8	24.4	1929	20	AAK31843	Nitrogen suboxid
5	27	23.7	534720	19	AAV30458	Rhizobium species
6	27	23.7	536165	19	AAV30459	Rhizobium species
7	26.8	23.5	286	19	AAV38059	Human stem cell a
8	26.6	23.3	1959	22	AA114146	Probe #4079 for g
9	26.6	23.3	1959	22	AA135531	Probe #4217 used
10	26.6	23.3	1959	22	AA103594	Probe #3985 used
11	26.2	23.0	236	22	AA150726	Probe #19412 used

C	45	24.8	21.8	948	22	AAI33938	Probe #6264 used t
C	44	24.8	21.8	948	22	AAI12590	Human inflammatory
C	43	24.8	21.8	700	22	AAH92837	Probe #523 for g
C	42	24.8	21.8	680	22	AAH90354	Human cDNA clone (
C	41	24.8	21.8	603	21	AAH90846	Hydrophobic domain
C	40	25	21.9	8148	19	AAV52144	Streptococcus pneu
C	39	25	21.9	1602	21	AAH50753	Arabidopsis thalial
C	38	25	21.9	1399	21	AAH39239	Arabidopsis thalial
C	37	25	21.9	1397	21	AAH50757	Arabidopsis thalial
C	36	25	21.9	1345	21	AAH39304	Arabidopsis thalial
C	35	25	21.9	954	18	AAT90219	Magnetospirillum s
C	34	25	21.9	740	21	AAH07055	Human secreted pro
C	33	25	21.9	315	19	AAH36954	Streptococcus pneu
C	32	25.2	22.1	49377	19	AAV05287	The sorphan biosy
C	31	25.2	22.1	28958	21	AAH5299	DNA sequence of So
C	30	25.2	22.1	28958	18	AAT69965	Sorangium cellulos
C	29	25.2	22.1	28598	17	AAH07679	Sorangium cellulos
C	28	25.2	22.1	3736	9	AAH84918	Rat brain specific
C	27	25.2	22.1	1438	9	AAH81177	Insert from pLpNK
C	26	25.2	22.1	932	16	AAH93079	Murine T-lif cDNA
C	25	25.2	22.3	813	20	AAH34155	Myobacterium spec
C	24	25.4	22.3	1134	20	AAH210576	DNA encoding g pro
C	23	25.4	22.3	1128	21	AAAH6026	Human g protein co
C	22	25.4	22.3	1128	21	AAAD01125	Human orphan G pr
C	21	25.4	22.3	1128	20	AAH210560	DNA encoding g pro
C	20	25.4	22.3	309	21	AAH54552	Human OREF ORF100
C	19	25.4	22.3	220	20	AAH34152	Myobacterium spec
C	18	25.4	22.3	219	20	AAH34153	Myobacterium spec
C	17	25.4	22.3	218	20	AAH34154	Myobacterium spec
C	16	25.6	22.6	659	20	AAH37662	Extended human se
C	15	25.8	22.6	5398	12	AAH33284	P.denitrificans co
C	14	25.8	22.6	1426	18	AAH96639	Human TUB Form 6
C	13	26	23.0	3129	14	AAH050349	VIP receptor . Rat
C	12	26.2	23.0	430	22	AAH37551	Probe #623/ used t

## ALIGNMENTS

RESULT	1
AAA91491	
ID	AAA91491 standard; DNA; 114 BP.
XX	
AC	AAA91491;
XX	
DT	06-AUG-2001 (first entry)
XX	
DE	G. oxydans cytochrome C oxidase (COIII) coding sequence fragment.
XX	
KW	Cytochrome C oxidase; COI; COII; COIII; COI complex; D-sorbitol;
KW	oxidative fermentation; electron transfer; respiratory chain; L-sorbose
KW	2-Keto-D-gluconic acid; 2KGa production; aldehyde production;
KW	carboxylic acid production; ketone production; ds.
OS	Glucanobacter oxydans.
XX	
Key	Location/Qualifiers
EH	1..114
FT	/*tag= a
FT	/product= "COIII"
XX	
PN	EP1103603-A2.
XX	
PD	30-MAY-2001.
XX	
PE	14-NOV-2000; 2000EP-0124785.
XX	
PR	17-NOV-1999; 99EP-0122842.
XX	
PA	(HOFF) HOFFMANN LA ROCHE & CO AG F.
XX	
PI	Asakura A, Hoshino T, Shinjoh M;
XX	

PR	04-NOV-1998;	98US-0172255.
PR	24-NOV-1998;	98US-0172252.
PR	22-DEC-1998;	98US-0172214.
PR	26-FEB-1999;	99US-0121896.
XX		
PA	(INCY-) INCYTE PHARM INC.	
XX		
PI	Hillman JL, Yue H, Tang YT, Lal P, Corley NC, Guegler KJ;	
XX	Baughn MK, Azimzai Y, Lu DAM,	
DR	P-PSTDB; AAY71067.	
XX		
PT	Novel human membrane transport proteins useful for diagnosis,	
PT	prevention and treatment of membrane transport disorders,	
PT	immune/inflammatory disorders and cell proliferative disorders	
XX	including cancer -	
XX		
PS	Claim 9; Page 119-120; 136pp; English.	
XX		
CC	The present sequence is a cDNA encoding membrane transport protein,	
CC	MTRP-12 from Incyte clone 1733477 isolated from human BRSTT08 cDNA	
CC	library. MTRP-12 shows homology to mouse transporter (GI 5359730)	
CC	and human transporter (GI 291964), and is expressed in nervous,	
CC	gastrointestinal and reproductive tissues.	
CC	The present sequence is useful in diagnosis, prevention and treatment	
CC	of disorders related with increased or decreased expression of MTRP	
CC	such as familial goitre, Menkes disease, diabetes, Parkinson's disease,	
CC	neurological disorders such as Alzheimer's disease, depression, epilepsy,	
CC	schizophrenia, immune/inflammatory disorders such as AIDS, Addison's	
CC	disease, allergies, atherosclerosis, Graves disease, gout, Hashimoto's	
CC	thyroiditis, viral, bacterial, fungal, parasitic, protozoal or	
CC	hematologic infections and cell proliferative disorders such as cancer.	
CC	Fragments of MTRP polynucleotides are useful as targets in microarrays.	
CC	MTRP DNA is also useful for generating hybridisation probes useful in	
CC	mapping genomic sequences and detecting differences in sequences among	
CC	normal, carrier and affected individuals. It is also used for	
CC	screening libraries of compounds in drug screening techniques.	
SQ	Sequence 2580 BP; 507 A; 766 C; 743 G; 564 T; 0 other:	
	Query Match	29.6%; Score 33.8; DB 21; Length 2580;
	Best Local Similarity	62.4%; Pred. No. 0.017;
	Matches 53; Conservative 0; Mismatches 32; Indels 0; Gaps 0;	
OY	2 tcgtccacggcagcccaagaanaaccgcgatctgacctagcgattgccgcgttgcct 61	
Db	1062 tctgttggaaggcgcttcacacalcgtgtgtgtgtgtgtgtgtgtgtgcactgcactcattgggga 1121	
OY	62 ggaactttaccctcgttcacaccta 86	
Db	1122 agatcttcgcctcgtgagcgacgcta 1146	
RESULF	3	
AAO70903/c		
ID	AAO70903 standard; cDNA; 1813 BP.	
XX		
AC	AAO70903;	
XX		
DT	20-FEB-1995 (first entry)	
XX		
DE	Cold acclimatization protein CAP85 cDNA from spinach leaf.	
XX		
KW	Cold acclimatization; cold-tolerance; transgenic plant; drought-	
KW	drought-resistance; drought-tolerance; crop improvement; CAP85;	
XX	freezing-tolerance; ss.	
OS	Spinacia oleracea.	
TH		
Key	Location/Qualifiers	
CDS	34..1641	
T		

```

FT  /tag= a
XX
XX WO9417186-A.
XX
XX 04-AUG-1994.
XX
XX 21-JAN-1994; 94WO-US00581.
XX
XX 21-JAN-1993; 93US-0007107.
XX
XX (UYFL ) UNIV FLORIDA.
XX
XX Guy CL, Haskell DW, Hoffig A, Neven LG;
XX
XX WPI; 1994-264100/32.
XX
XX P-PSDB: AAR56549.
XX
XX Cold acclimatization proteins CAP85 and CAP160 and their
XX nucleotide sequences - used to confer increased cold tolerance
XX and drought resistance on plants and microorganisms by genetic
XX transformation
XX
XX Claim 4; Page 28; 49pp; English.
XX
XX Transgenic plants constructed using this cold acclimatization
XX protein nucleotide sequence have improved cold tolerance and/or
XX drought-resistance. The sequence can also be used as a probe in
XX assays for crop and plant tolerance levels during seasons of risk to
XX freezing temp. or drought conditions. The cell transformed with
XX this sequence is from the family Solanaceae, a citrus plant, a
XX bacterium or a yeast cell.
XX
XX Sequence 1813 BP; 645 A; 392 C; 409 G; 367 T; 0 other:
XX

```

```

Query Match          27.2%; Score 31; DB 15; Length 1813;
Best Local Similarity 56.3%; Pred. No. 0.16;
Matches 58; Conservative 0; Mismatches 45; Indels 0; Gaps 0;

```

```

QY 9 cggcgacgcgaagaacgcgattgctagcattgcacatgccttgctgctatctt 68
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
DB 196 GCGAGACGGGCACTTCCGATACCGCGAGCGCATCATCGGCTCGATGCTTGT 137
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 69 taccctgtgccaaagcctatgaatatatgaatcgtccatacc 111
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
DB 136 GATCCTCTCCCTTGTGTTCTTCTTCATGAATCAACATACCC 94
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||

```

## RESULT 4

AAK31843  
ID AAK31843 standard; DNA; 1929 BP.

AAK31843;

14-JUN-1999 (first entry)

Nitrogen suboxide reductase encoding DNA.

Nitrogen suboxide reductase; transgenic plant; fortified absorption;  
decomposition; nitrogen oxide; ds.

Achromobacter cycloclastes.

JF11075842-A.

23-MAR-1999.

29-AUG-1997; 97JP-0234200.

29-AUG-1997; 97JP-0234200.

(MITO ) MITSUBISHI ELECTRIC CORP.

XX

```

DR WPI; 1999-257693/22.
XX
XX P-PSDB; AAY03775.
XX
XX New nitrous oxide reductase gene and protein - useful for production
XX of transgenic plants with fortified absorption and decomposition of
XX nitrogen oxide
XX
XX Claim 2; Pages 7-9; 10pp; Japanese.
XX
XX This represents a A. cycloclastes nitrogen suboxide reductase encoding
XX gene. The gene can be used for the production of transgenic plants with
XX fortified absorption. A microbe and a plant of fortified absorption and
XX decomposition of nitrogen oxide can be produced by recombining the gene
XX to the cell of the microbe and the plant.
XX
XX Sequence 1929 BP; 427 A; 593 C; 578 G; 331 T; 0 other:
XX

```

```

Query Match          24.4%; Score 27.8; DB 20; Length 1929;
Best Local Similarity 59.5%; Pred. No. 2.4;
Matches 47; Conservative 0; Mismatches 32; Indels 0; Gaps 0;

```

```

QY 7 caccgacgcgaagaacgcgattgctagcattgcacatgccttgctgctatctt 66
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
DB 1288 caaccgacactgaagacggtgattgagcgaacgctgattgcgcacacgactgctc 1347
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 67 ttaccctgtgccaaagcct 85
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
DB 1348 gtttgctgtgcaaatctt 1366
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||

```

## RESULT 5

AAV30458  
ID AAV30458 standard; DNA; 534720 BP.

AAV30458;

14-OCT-1998 (first entry)

Rhizobium species plasmid pNGR234a.

Symbiosis; open reading frame; ORF; plasmid; vector; transportation;  
degradation; metabolism; host range; nitrogen fixation; nodulation;

legume; plant; ds.

Rhizobium sp.

Location/Qualifiers

417796..418671

/tag= a

/standard\_name= "ORF K1"

/note= "oligopeptide permease"

418673..419580

/tag= b

/standard\_name= "ORF K2"

/product= "oligopeptide permease"

419677..420738

/note= "homologous to the oppd gene"

420774..422159

/tag= d

/standard\_name= "ORF K4"

/product= "encapsulation-like protein"

/note= "homologous to the CapA gene"

422628..424031

/tag= e

/standard\_name= "ORF K5"

/product= "aminotransferase-like protein"

/note= "homologous to the BioA gene"

XX

```

FT CDS 424056..425594
FT /tag= f
FT /standard_name= "ORF K6"
FT /product= "(semi)aldehyde dehydrogenase-like protein"
FT complement (426949..428028)
FT CDS
FT /tag= g
FT /standard_name= "ORF K7"
FT /product= "transposase homologue"
FT /note= "homologous to the Trp gene"
FT 428292..429623
FT /tag= h
FT /standard_name= "ORF K8"
FT /product= "glutamate dehydrogenase-like protein"
FT /note= "homologous to the GluT1 gene"
FT complement (430538..431284)
FT CDS
FT /tag= i
FT /standard_name= "ORF K9"
FT /product= "transposase homologue"
FT complement (431296..432840)
FT CDS
FT /tag= j
FT /standard_name= "ORF K10"
FT /product= "transposase homologue"
FT /note= "homologous to the Trp gene"
FT complement (433880..434110)
FT CDS
FT /tag= k
FT /standard_name= "ORF K11"
FT /product= "protein of unknown function"
FT /note= "homologous to the fixu gene"
FT complement (434107..434433)
FT CDS
FT /tag= l
FT /standard_name= "ORF K12"
FT /product= "protein of unknown function"
FT complement (434517..434711)
FT CDS
FT /tag= m
FT /standard_name= "ORF K13"
FT /product= "ferrodoxin/ferrodoxin-like protein"
FT /note= "homologous to the FdxN gene"
FT complement (434753..436234)
FT CDS
FT /tag= n
FT /standard_name= "ORF K14"
FT /gene= "nifB"
FT /product= "protein involved in Femo co-factor biosynthesis"
FT complement (436460..438130)
FT CDS
FT /tag= o
FT /standard_name= "ORF K15"
FT /gene= "nifA"
FT /product= "positive regulator of nif, fix and other genes"
FT complement (438297..438590)
FT CDS
FT /tag= p
FT /standard_name= "ORF K16"
FT /gene= "fixX"
FT /product= "protein required for nitrogenase activity"
FT complement (438605..439912)
FT CDS
FT /tag= q
FT /standard_name= "ORF K17"
FT /gene= "fixC"
FT /product= "protein required for nitrogenase activity"
FT complement (439923..441032)
FT CDS
FT /tag= r
FT /standard_name= "ORF K18"
FT /gene= "fixB"
FT /product= "protein required for nitrogenase activity"
FT complement (441042..441899)
FT CDS
FT /tag= s
FT /standard_name= "ORF K19"
FT /gene= "fixA"
FT /product= "protein required for nitrogenase activity"
FT complement (442316..442636)
FT CDS
FT /tag= t
FT /standard_name= "ORF K20"
FT /product= "protein of unknown function"

```

---

```

FT CDS complement (443313..443879)
FT /tag= u
FT /standard_name= "ORF K21"
FT /product= "protein of unknown function"
FT 444337..445029
FT CDS
FT /tag= v
FT /standard_name= "ORF K22"
FT /product= "ferrodoxin-like protein"
FT /note= "homologous to the NifQ gene"
FT 445088..446602
FT CDS
FT /tag= w
FT /standard_name= "ORF K23"
FT /gene= "dctA"
FT /product= "C4-dicarboxylate transport protein"
FT /note= "homologous to the DctAI gene"
FT 446599..447843
FT CDS
FT /tag= x
FT /standard_name= "ORF L1"
FT /product= "cytochrome P450-like protein"
FT /note= "homologous to the CamC gene"
FT 447844..448500
FT CDS
FT /tag= y
FT /standard_name= "ORF L2"
FT /product= "gamma-hexachlorocyclohexane-dechlorinase-like protein"
FT 448497..450203
FT CDS
FT /tag= z
FT /standard_name= "ORF L3"
FT /product= "putative protein with degradative function"
FT 450341..451396
FT CDS
FT /tag= aa
FT /standard_name= "ORF L4"
FT /product= "luciferase alpha-subunit-like protein"
FT /note= "homologous to the LuxA gene"
FT 452980..454494
FT CDS
FT /tag= ab
FT /standard_name= "ORF L6"
FT /gene= "nifD"
FT /product= "alpha-subunit of Femo protein of nitrogenase"
FT 454590..456131
FT CDS
FT /tag= ac
FT /standard_name= "ORF L7"
FT /gene= "nifK"
FT /product= "beta-subunit of Femo protein of nitrogenase"
FT 456187..457677
FT CDS
FT /tag= ad
FT /standard_name= "ORF L8"
FT /product= "protein involved in Femo co-factor biosynthesis"
FT /note= "homologous to the NifB gene"
FT 457687..459096
FT CDS
FT /tag= ae
FT /standard_name= "ORF L9"
FT /product= "protein involved in Femo co-factor biosynthesis"
FT /note= "homologous to the fixF gene"
FT 459093..459575
FT CDS
FT /tag= af
FT /standard_name= "ORF L10"
FT /product= "protein of unknown function"
FT /note= "homologous to the NifX gene"
FT 459579..460067
FT CDS
FT /tag= ag
FT /standard_name= "ORF L11"
FT /product= "protein of unknown function"
FT 460501..460920
FT CDS
FT /tag= ah
FT /standard_name= "ORF L12"
FT /product= "protein similar to part of the Fe protein of nitrogenase"
FT /note= "homologous to the NifH gene"
FT 461228..461545

```

```

FT      /tag= a1
FT      /standard_name= "ORF L13"
FT      /product= "protein of unknown function"
FT      463201..464733
FT      CDS
FT      /tag= a1
FT      /standard_name= "ORF L14"
FT      /product= "peptidase-like protein"
FT      /note= "homologous to the bi-MP gene"
FT      464736..466079
FT      CDS
FT      /tag= ak
FT      /standard_name= "ORF L15"
FT      /product= "processing protease-like protein"
FT      /note= "homologous to the PP gene"
FT      466590..467021
FT      CDS

Query Match      23.7%; Score 27; DB 19; Length 534720;
Best Local Similarity 66.1%; Pred. No. 32;
Matches 39; Conservative 0; Mismatches 20; Indels 0; Gaps 0;

QY      2 tcgtccacgcgcacgcagaacccgcgattgctgctagcattgcacatcgcccttgc 60
        |||| |||||||||||| ||| || ||||| || || || ||
DB 189310 tcgttaacgcgcacgcacgaatgcgcggagccttgagagcggaacgcactccgc 189368

RESULT 6
AAV30459
ID      AAV30459 standard; DNA; 536165 BP.
XX
AC      AAV30459;
XX
DT      06-JUL-1999 (first entry)
XX
DE      Rhizobium species symbiotic plasmid pNGR234.
XX
KW      Symbiosis; open reading frame; ORF; plasmid; vector; transportation;
KW      degradation; metabolism; host range; nitrogen fixation; nodulation;
KW      legume; plant; ds.
XX
XX      Rhizobium sp.
XX      OS
XX      WO9802560-A2.
XX      PN
XX      22-JAN-1998.
XX      PD
XX      10-JUL-1997; 97WO-IB00950.
XX      PE
XX      20-MAY-1997; 97GB-0010395.
XX      PR
XX      12-JUL-1996; 96EP-0730001.
XX
PA      (MOLE-) INST MOLECULAR BIOTECHNOLOGY.
PA      (BIOL-) LAB BIOLOGIE MOLECULAIRE PLANTES SUPRIE.
XX
PI      Broughton WJ, Freiberg CB, Perret XP, Rosenthal A;
XX
DR      WPI, 1998-110606/10.
XX
PT      New isolated symbiotic plasmid from Rhizobium sp. NGR234 - used to
PT      develop products for modifying plant characteristics; e.g. nitrogen
PT      fixation, synthesis of compounds and stress response
XX
XX      Claim 1: Fig 3; 228pp; English.
XX
XX      This is the nucleotide sequence of the plasmid pNGR234a isolated from
XX      Rhizobium sp. NGR234. Open reading frames (ORF) derivable from the
XX      nucleotide sequence are claimed. The nucleotide sequences or ORFs can
XX      be used e.g. in the transportation of compounds to and from an organism
XX      which is a host to at least one of the nucleotide sequences, ORFs or
XX      proteins, the degradation and/or metabolism of organic, inorganic,
XX      natural or xenobiotic substances in a host organism or the modification
XX      of the host range, nitrogen fixation abilities; for obtaining a synthetic
XX      minimal set of ORFs required for functional Rhizobium-legume symbiosis,
XX      especially for nodulation efficiency on host plants.
XX

```

```

SQ      Sequence 536165 BP; 111291 A; 155755 C; 157864 G; 111255 T; 0 other;

Query Match      23.7%; Score 27; DB 19; Length 536165;
Best Local Similarity 66.1%; Pred. No. 32;
Matches 39; Conservative 0; Mismatches 20; Indels 0; Gaps 0;

QY      2 tcgtccacgcgcacgcagaacccgcgattgctgctagcattgcacatcgcccttgc 60
        |||| |||||||||||| ||| || ||||| || || || ||
DB 189310 tcgttaacgcgcacgcacgaatgcgcggagccttgagagcggaacgcactccgc 189368

RESULT 7
AAV38059
ID      AAV38059 standard; CDNA; 286 BP.
XX
AC      AAV38059;
XX
DT      14-SEP-1998 (first entry)
XX
DE      Human stem cell antigen 2 partial sequence from Incyte clone 588615.
XX
XX      Human; stem cell antigen; SCAH-1; SCAH-2; THP-1; bladder tumour;
XX      diagnosis; screening; sca-2; ly-6 family; leukaemia; cancer; ss.
XX
XX      Homo sapiens.
XX      OS
XX      WO9800540-A1.
XX      PN
XX      08-JAN-1998.
XX      PD
XX      25-JUN-1997; 97WO-US10956.
XX      PE
XX      03-JUL-1996; 96US-0675508.
XX
PA      (INCY-) INCYTE PHARM INC.
XX
PI      Au-Young J;
XX
DR      WPI, 1998-427478/36.
XX
PT      Human stem cell antigens, SCAH-1 and SCAH-2 - useful to, e.g.
PT      diagnose and treat SCAH-1 or SCAH-2 related conditions such as
PT      tumours and screen inhibitory compounds
XX
XX      Disclosure; Page 49; 66pp; English.
XX
XX      The present sequence represents an Incyte clone used in the determination
XX      of human stem cell antigen 2 (SCAH-2). SCAH proteins can be administered
XX      therapeutically, e.g. to inhibit or reverse the development of tumours.
XX      Antibodies to SCAH proteins can be used in diagnostic tests for
XX      conditions associated with protein expression in biological samples, by
XX      combining a sample with the antibody under conditions suitable for
XX      antibody binding to the protein to form a complex, and detecting the
XX      complex. Antibodies to SCAH-2 are especially useful in diagnostic tests
XX      for conditions/diseases such as leukaemias or malignant local tumours
XX      associated with SCAH-2 expression. SCAH-1 and SCAH-2 have homology to
XX      SCA-2 and other stem cell antigens, and have characteristics of the
XX      ly-6 family of cysteine rich proteins. The functional similarities among
XX      ly-6 family proteins previously reported, and the expression of SCAH-1
XX      and SCAH-2 in tumours from several tissues (e.g. breast, lung, bladder)
XX      indicated that SCAH proteins may be useful as anticancer agents; SCAH
XX      antibodies, antagonists or inhibitors identified using the proteins
XX      could also be useful in anticancer treatments and to intervene in
XX      allorresponses associated with transplant rejection and autoimmune
XX      diseases e.g. lupus nephritis. The nucleic acid sequences encoding SCAH
XX      proteins are also useful therapeutically to treat, e.g. leukaemias and
XX      cancers of the bladder or breast.
XX
SQ      Sequence 286 BP; 58 A; 82 C; 90 G; 47 T; 9 other;

Query Match      23.5%; Score 26.8; DB 19; Length 286;

```





XX	09-OCT-2001	(first entry)
DT		
XX		
XX	Probe #3985	used to measure gene expression in human breast sample.
XX		
XX	Probe;	human; breast disease; breast cancer; development disorder; ss;
KW	Inflammatory	disease; proliferative breast disease; non-carcinoma tumour.
XX		
XX	Homo sapiens.	
XX		
XX	MO200157270-A2.	
PN		
PD		
XX	09-AUG-2001.	
XX		
XX	29-JAN-2001;	2001WO-US00661.
PE		
XX		
XX	04-FEB-2000;	2000US-0180312.
PR	26-MAY-2000;	2000US-0207456.
PR	30-JUN-2000;	2000US-0608408.
PR	03-AUG-2000;	2000US-0632366.
PR	21-SEP-2000;	2000US-0234687.
PR	27-SEP-2000;	2000US-0236359.
PR	04-OCT-2000;	2000GB-0024263.
XX		
PA	(MOLE-)	MOLECULAR DYNAMICS INC.
XX		
XX	Penn SG,	Hanzel DK, Chen W, Rank DR;
XX		
XX	WPI,	2001-476286/51.
DR		
XX		
PT	Novel single exon	nucleic acid probe used to measuring gene expression
PT	in a human breast	-
PS		
XX	Claim 25;	SEQ ID No 3985; 322pp; English.
XX		
CC	The present invention	relates to novel single exon nucleic acid probes.
CC	The present sequence	is one such probe. The probes are useful for
CC	measuring human gene	expression in a human breast sample, where the probe
CC	hybridises at high	stringency to a nucleic acid expressed in the human
CC	breast. The probes	are useful for predicting, diagnosing, grading,
CC	staging, monitoring	and prognosing diseases of the human breast,
CC	particularly those	diseases with polygenic aetiology. The diseases
CC	include: breast	cancer, disorders of development, inflammatory diseases
CC	of the breast,	fibrocystic changes, proliferative breast disease and
CC	non-carcinoma	tumours.
CC	Note:	The sequence data for this patent did not form part of the printed
CC	specification,	but was obtained in electronic format directly from Wipo
CC	at	ftp.wipo.int/pub/published_pct-sequences.
XX		
XX	Sequence	1959 BP; 521 A; 343 C; 564 G; 531 T; 0 other;
SO		
	Query Match	23.3%; Score 26.6; DB 22; Length 1959;
	Best Local Similarity	54.6%; Pred. No. 6.5;
	Matches 53; Conservative	0; Mismatches 44; Indels 0; Gaps 0;
Qy	1	atcgttcacgagcgacgcaagaaacgcgattggcctagcagatgcacatcgccctggc 60
Db	443	ATCATCAACCCATCTCATCTACCACTACCGCTCAGGCTATCATCATCAACCACTATGTT 384
Qy	61	tgagatcttiacccttgycacagcctatgaataatag 97
Db	383	GTCACTCATCATCTCTCTCAAGCCTATGACCTTTTG 347
RESULT 11		
AA150726		
ID	AA150726	standard; DNA; 236 BP.
XX		
AC	AA150726;	
XX		
DT	17-OCT-2001	(first entry)
XX		

DE		Probe #19412 used to measure gene expression in human placenta sample.
XX	KW	Probe; microarray; human; placenta; antenatal diagnosis;
XX	KW	genetic disorder; ss.
OS	Homo sapiens .	
XX	PN	WO200157272-A2.
PD	XX	
XX	PD	09-AUG-2001.
PF	XX	
PE	XX	30-JAN-2001; 2001WO-US00663.
XX	PR	04-FEB-2000; 2000US-0180312.
PR	PR	26-MAY-2000; 2000US-0207456.
PR	PR	30-JUN-2000; 2000US-0608408.
PR	PR	03-AUG-2000; 2000US-0632366.
PR	PR	21-SEP-2000; 2000US-0234687.
PR	PR	27-SEP-2000; 2000US-0236359.
PR	PR	04-OCT-2000; 2000GB-0024263.
XX		
PA	(MOLE-) MOLECULAR DYNAMICS INC.	
XX	PI	Penn SG, Hanzel DK, Chen W, Rank DR;
XX	DR	WPI; 2001-488897/53.
XX	PT	Human genome-derived single exon nucleic acid probes useful for analyzing gene expression in human placenta -
PS	Claim 25; SEQ ID No 19412; 654bp; English.	
CC	The present invention relates to single exon nucleic acid probes (SENPs).	
CC	The present sequence is one such probe. The probes are useful for producing a microarray for predicting, measuring and displaying gene expression in samples derived from human placenta. The probes are useful for antenatal diagnosis of human genetic disorders.	
CC		
SQ	Sequence 236 BP; 88 A; 60 C; 37 G; 51 T; 0 other:	
	Query Match	23.0%; Score 26.2; DB 22; Length 236;
	Best Local Similarity	53.4%; Pred. No. 4.4;
	Matches	55; Conservative 0; Mismatches 48; Indels 0; Gaps 0;
QY	6 ccaagcgacccgaagaaccgcgattggcctagcgttcgccatcgcccttgcgtgat 65 	
DB	79 ccacgagcacccgcacaagaacctttacaaggccttgatgtattttccaccacc 138 	
QY	66 ctllaccctgtgccaaagcctaatgaatatgaatcgtccaat 108 	
DB	139 ctltgtactgtlccataaccttaaaaatacaagaagaacaactat 181 	
	RESULT 12	
ID	AAI37551	
XX	AAI37551 standard; DNA; 430 BP.	
AC	AAI37551;	
XX		
DT	17-OCT-2001 (first entry)	
DE	Probe #6237 used to measure gene expression in human placenta sample.	
XX	Probe; microarray; human; placenta; antenatal diagnosis;	
KM	genetic disorder; ss.	
XX		
OS	Homo sapiens.	
XX		
PN	WO200157272-A2.	
PD	09-AUG-2001.	
XX		

PF 30-JAN-2001; 2001WO-US00663.  
 XX  
 PR 04-FEB-2000; 2000US-0180312.  
 PR 26-MAY-2000; 2000US-0207456.  
 PR 30-JUN-2000; 2000US-0608408.  
 PR 03-AUG-2000; 2000US-0632366.  
 PR 21-SEP-2000; 2000US-0234687.  
 PR 27-SEP-2000; 2000US-0236359.  
 PR 04-OCT-2000; 2000US-0024263.  
 XX

(MOLE-) MOLECULAR DYNAMICS INC.

PI Penn SG, Hanzel DK, Chen W, Rank DR;  
 DR WPI; 2001-488897/53.  
 XX

PT Human genome-derived single exon nucleic acid probes useful for  
 XX analyzing gene expression in human placenta -  
 XX

PS Claim 25; SEQ ID NO 6237; 654bp; English.

CC The present invention relates to single exon nucleic acid probes (SENP).  
 CC The present sequence is one such probe. The probes are useful for  
 CC producing a microarray for predicting, measuring and displaying gene  
 CC expression in samples derived from human placenta. The probes are useful  
 CC for antenatal diagnosis of human genetic disorders.  
 XX

Sequence 430 BP; 154 A; 96 C; 61 G; 119 T; 0 other;

Query Match 23.0%; Score 26.2; DB 22; Length 430;  
 Best Local Similarity 53.4%; Pred. No. 5.4;  
 Matches 55; Conservative 0; Mismatches 48; Indels 0; Gaps 0;

OY 6 ccacggcgacgcgaagaacacgcgattgctcagcattgcgcctgctgctgcat 65  
 DB 293 ccacggcgacgcgaagaacacgccttacaagaagccttgatctcttaccaccac 352  
 OY 66 cttaaccctgtgccaagcctatgaatatattgaatgctcat 108  
 DB 353 ctgtgactgtgctcattacataaatacaagaagaactact 395

RESULT 13

AAQ50349  
 ID AAQ50349 standard; DNA: 3129 BP.

AC AAQ50349;

DT 13-MAY-1994 (first entry)

DE VIP receptor.

KM vasoactive intestinal polypeptide receptor; VIP; rat; binding;  
 KM adenylate cyclase activity; stimulus; ds.  
 XX

OS Rattus rattus.

Key Location/Qualifiers  
 FT CDS 59..1438  
 FT /tag= a  
 FT /product= VIP\_receptor

PN JP05355394-A.

PD 05-OCT-1993.

PF 13-FEB-1992; 92JP-0026607.

PR 13-FEB-1992; 92JP-0026607.

PA (OSAB-) ZH OSAKA BIOSCIENCE KENKYUSHO.  
 XX

DR WPI: 1993-348480/44.  
 DR P-PSDB; AAR42848.

PT Vasoactive intestinal polypeptide - prepd. in large amt. by  
 XX culturing microbe transformed by new DNA coding polypeptide  
 XX  
 PS Claim 4; Page 6; 14pp; Japanese.

CC The sequence can be used to produce large amounts of the VIP  
 CC receptor peptide, by culturing a microorganism transformed by  
 CC the sequence.  
 XX

Sequence 3129 BP; 671 A; 915 C; 852 G; 691 T; 0 other;

Query Match 22.8%; Score 26; DB 14; Length 3129;  
 Best Local Similarity 52.8%; Pred. No. 13;  
 Matches 56; Conservative 0; Mismatches 50; Indels 0; Gaps 0;

OY 6 ccacggcgacgcgaagaacacgcgattgctcagcattgcgcctgctgctgcat 65  
 DB 265 cctcaccctgtgcccgaacaaccccgaggcgagcgtgctgctgctgctgctcctcat 324  
 OY 66 cttaaccctgtgccaagcctatgaatatattgaatgctcatacc 111  
 DB 325 cttaaccctgtgcccattcattgattataacatacagcgcgtgac 370

RESULT 14

AAT96639  
 ID AAT96639 standard; cDNA: 1426 BP.

AC AAT96639;

DT 27-APR-1998 (first entry)

DE Human TUB Form 6 cDNA.

KM TUB Form 6; sensory neuron; neurosensory defect;  
 KW cochlear degeneration; hearing loss; deafness; retinal dystrophy;  
 KW retinitis pigmentosa; combined rod cone dystrophy; obesity; human;  
 KW animal model; transgenic animal; therapy; diagnosis; ds.  
 XX

OS Homo sapiens.

Key Location/Qualifiers  
 FT CDS 28..1410  
 FT /tag= a

PN WO9738004-A1.

PD 16-OCT-1997.

PF 10-APR-1997; 97WO-US05903.

PR 17-SEP-1996; 96US-0714991.

PR 10-APR-1996; 96US-0630592.

PR 22-AUG-1996; 96US-0701380.

PR 04-SEP-1996; 96US-0706292.

PA (JACK-) JACKSON LAB.

PI (SEOU-) SEOULANA THERAPEUTICS INC.

PI Nagert J, Nishina P, Noben-Trauth K, North M;

DR WPI: 1997-512642/47.

DR P-PSDB; AAM36488.  
 XX  
 PT Mammalian TUB protein - used for detecting pre-disposition to  
 PT neuro-sensory defects  
 XX  
 PS Disclosure: Page 48-49; 89pp; English.  
 XX



Thu Jan 10 08:15:41 2002

us-09-712-768-5.rng

---

Page 10

GenCore version 4.5  
Copyright (c) 1993 - 2000 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: January 10, 2002, 00:05:12 ; Search time 1374.03 Seconds  
(without alignments)  
1368.732 Million cell updates/sec

Title: US-09-712-768-5  
Sequence: 1 atcgccacggcgaccgca.....atgaatcgtccatccgca 114

Scoring table: IDENTITY\_NDC  
Gapop 10.0 , Gapext 1.0

Searched: 1472140 seqs, 8248589755 residues  
Total number of hits satisfying chosen parameters: 2944280

Minimum DB seq length: 0  
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%  
Maximum Match 100%  
Listing first 45 summaries

Database :

GenEmbl: \*  
1: gb\_ba: \*  
2: gb\_htg: \*  
3: gb\_in: \*  
4: gb\_cm: \*  
5: gb\_ov: \*  
6: gb\_pat: \*  
7: gb\_ph: \*  
8: gb\_pl: \*  
9: gb\_pr: \*  
10: gb\_ro: \*  
11: gb\_sts: \*  
12: gb\_sy: \*  
13: gb\_un: \*  
14: gb\_vl: \*  
15: em\_ba: \*  
16: em\_fun: \*  
17: em\_hum: \*  
18: em\_in: \*  
19: em\_om: \*  
20: em\_or: \*  
21: em\_ov: \*  
22: em\_pat: \*  
23: em\_ph: \*  
24: em\_pl: \*  
25: em\_ro: \*  
26: em\_sts: \*  
27: em\_sy: \*  
28: em\_un: \*  
29: em\_vl: \*  
30: em\_htgo\_hum: \*  
31: em\_htgo\_inv: \*  
32: em\_htgo\_rnd: \*  
33: em\_htg\_hum: \*  
34: em\_htg\_inv: \*  
35: em\_htg\_rnd: \*  
36: em\_htg\_other: \*

pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

## SUMMARIES

Result No.	Score	Query Length	DB ID	Description
1	114	100.0	114 6 AX113990	AX113990 Sequence
2	39.8	34.9	346510 1 AP003011	AP003011 Mesorhizobium loti
3	37.6	33.0	17056 4 MTDCNOMGN	MTDCNOMGN
4	37	32.5	16507 10 SVU238588	SVU238588
5	37	32.5	16968 4 AF034071	AF034071
6	35.6	31.2	783 4 AF034071	AF034071
7	35.6	31.2	1130 4 CA085910	CA085910
8	35.6	31.2	16616 4 AF010406	AF010406
9	35.4	31.1	2215 4 RABAIAR	RABAIAR
10	35.4	31.1	16428 4 PMAT277029	PMAT277029
11	35.4	31.1	16826 4 MIPVNA	MIPVNA
12	35.4	31.1	17245 4 AF030461	AF030461
13	35.2	30.9	783 4 AF030472	AF030472
14	34.8	30.5	783 4 AF030473	AF030473
15	34.8	30.5	783 4 AF030470	AF030470
16	34.4	30.2	783 4 AF0304202	AF0304202
17	34.4	30.2	15977 4 AF0304203	AF0304203
18	34.4	30.2	15978 4 AF0304200	AF0304200
19	34.4	30.2	15978 4 AF0304253	AF0304253
20	34.4	30.2	16613 4 SS002189	SS002189
21	34.4	30.2	16680 4 AF030489	AF030489
22	34.2	30.0	17804 5 AF288410	AF288410
23	33.8	29.6	2589 9 AF279265	AF279265
24	33.8	29.6	2748 9 AC005923	AC005923
25	33.8	29.6	88326 2 AC024179	AC024179
26	33.8	29.6	204513 2 AF030464	AF030464
27	33.2	29.1	783 4 AF030482	AF030482
28	33.2	29.1	783 4 AF030483	AF030483
29	33.2	29.1	783 4 AF030484	AF030484
30	33.2	29.1	783 4 AF030271	AF030271
31	32.8	28.8	783 4 AF030275	AF030275
32	32.8	28.8	783 4 AF030278	AF030278
33	32.8	28.8	783 4 AF030460	AF030460
34	32.8	28.8	783 4 AF030469	AF030469
35	32.8	28.8	783 4 AF030457	AF030457
36	32.8	28.8	15979 4 AF0304201	AF0304201
37	32.8	28.8	16660 4 MIECCOMP	MIECCOMP
38	32.4	28.4	783 4 AF030457	AF030457
39	32.4	28.4	783 4 AF030457	AF030457
40	32.4	28.4	783 4 AF030474	AF030474
41	32.2	28.2	16651 4 AF061340	AF061340
42	32.2	28.2	202859 2 AC073547	AC073547
43	32	28.1	783 4 AF030274	AF030274
44	32	28.1	783 4 AF030479	AF030479
45	32	28.1	783 4 AF030481	AF030481

## ALIGNMENTS

RESULT 1  
AX113990 114 bp DNA PAT 01-MAY-2001  
LOCUS AX113990 Sequence 5 from Patent EP1103603.  
DEFINITION AX113990  
ACCESSION AX113990  
VERSION AX113990.1 GI:13940150  
KEYWORDS  
SOURCE  
ORGANISM  
Glucobacter oxydans.  
Bacteria; Proteobacteria; alpha subdivision; Acetobacteraceae;

REFERENCE  
1 (bases 1 to 114)  
Asakura, A., Hoshino, T. and Shinoh, M.  
Cytochrome c oxidase complex from glucobacter oxydans  
Patent: EP 1103603-A 5 30-MAY-2001;  
F. HOFFMANN-LA ROCHE AG (CH)  
JOURNAL  
Location/Qualifiers  
FEATURES  
source  
1..114  
/organism="Glucobacter oxydans"  
/db\_xref="taxon:442"

```

CDS
1..>114
/note="unnamed protein product"
/codon_start=1
/transl_table=11
/protein_id="CAC37871.1"
/db_xref="GI:13940151"
/translation="MVHGRKRTAIGLAIAIGLWITLCAVEYEVIVHTE"

BASE COUNT
29 a 33 c 26 g 26 t

Query Match
Best Local Similarity 100.0%; Score 114; DB 6; Length 114;
Matches 114; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 atcgtccacggcgaccgaagaacacgcgattggcctaagcattgcacgcctggc 60
Db 1 ATCGTCCACGGCGACCGACGAAGAAACCGCATGCGCATGCGCATGCGCATGCGC 60

Qy 61 tggattcttacccttgctcgaagcctatgaatattgaatgaatcgtccatccgaa 114
Db 61 TGGATCTTACCCTGTGCGCAGCCTATGATATTATGAATCGTCATACCGAA 114

RESULT 2
AP003011 346510 bp DNA BCT 15-MAY-2001
DEFINITION Mesorhizobium loti DNA, complete genome, section 18/21.
ACCESSION AP003011 BA000012
VERSION AP003011.2 GI:14026664
KEYWORDS
SOURCE Mesorhizobium loti (strain:MAFP303099) DNA.
ORGANISM Mesorhizobium loti
Bacteria: Proteobacteria; alpha subdivision: Rhizobiaceae group;
Phyllobacteriaceae; Mesorhizobium.
1 (sites)
Kaneko, T., Nakamura, Y., Sato, S., Asamizu, E., Kato, T., Sasamoto, S.,
Watanabe, A., Idesawa, K., Ishikawa, A., Kawashima, K., Kimura, T.,
Kishida, Y., Kiyokawa, C., Kohara, M., Matsumoto, M., Matsuno, A.,
Mochizuki, Y., Nakayama, S., Nakazaki, N., Shimpo, S., Sugimoto, M.,
Takeuchi, C., Yamada, M. and Tabata, S.
Complete genome structure of the nitrogen-fixing symbiotic
bacterium Mesorhizobium loti
DNA Res. 7 (6), 331-338 (2000)
2 (bases 1 to 346510)
Kaneko, T.
Direct Submission
Submitted (05-DEC-2000) Takakazu Kaneko, Kazusa DNA Research
Institute, The First Laboratory for Plant Gene Research, Yana
1532-3, Kisarazu, Chiba 292-0812, Japan
(E-mail:kaneko@kazusa.or.jp,
URL:http://www.kazusa.or.jp/rhizobae/,
Tel:81-438-52-3935(ex.2338), Fax:81-438-52-3934)
On May 11, 2001 this sequence version replaced gi:11994966.
Location/Qualifiers
1. 346510
/organism="Mesorhizobium loti"
/strain="MAFP303099"
/db_xref="taxon:381"
/complement(242..1264)
/gene="mlr7087"
/complement(242..1264)
/feature="note" similar to lipopolysaccharide core biosynthesis
glycosyl transferase Lpse"
/codon_start=1
/transl_table=11
/protein_id="BAB53261.1"
/db_xref="GI:14026665"
/translation="MAGALACAVLRVSVHPRILAIADRPMAIDKTKONTLRVLER
RFEIVOROHVTSVDSVNAADIVITFWQIAKLPESVLRSDRLVIGICSHREU
EGEPRERELATLHRLRAVAVFVNNRKLHEYAPLRIRIPVHYTPNGVDITTEFCRPREOP
YSGELRVAGSLAGNHGPAHGFHDVIEPANAIPGARILTAIREQMRNHDMLDFY
RDLIVYVCAASRESETPNCLPAACGIFLVTTRVGNMDELVOGDNLIFEDGAEGLA
DKALLRDTPTSLRSMAARMTETLRGDMRWQVQANVHMTSLTGGAAVEPRRAAS
DVRGOT"
/complement(1228..2265)
/gene="mlr7088"
/complement(1228..2265)
/gene="mlr7088"
/feature="note" similar to lipopolysaccharide core biosynthesis
glycosyl transferase Lpse"
/codon_start=1
/transl_table=11
/protein_id="BAB53262.1"
/db_xref="GI:14026666"
/translation="MALINGEHLIDLOGGVSTKCGALAMPSSLIRLVISAKFSNCALV
ERPAMHEAVNBEGRNVELNTHIAHAERNSPRVAVLTTRNEERFVACLEULFH
OGVOVYCDNOSDRITLQRTYRGAGLIGFESIPHNGWYCEQOLLRKEELFOSLEA
DWLMLHLDRIHLPYISYSLVALAADALGPNAIESESEFTEITREAPHONDPO
HTLRITTYPERFSPHCVRAYKKODGPIEIAMGGHVRVREGPVRLSPRPRKHTFL
SABHAARKYGGRRFLESEVDGRGWHGMRPLRPEDIQLPDITQVTTATYDDLDDEANP
WLAHWLARVAS"
/complement(2648..3787)
/gene="mlr7089"
/complement(2648..3787)
/gene="mlr7089"
/feature="note" similar to mandelate racemase"
/codon_start=1
/transl_table=11
/protein_id="BAB53263.1"
/db_xref="GI:14026667"
/translation="MSLSHFRTITRFQFQRDRVIGDSQVADNVVAALVELSEGEV
IGFIQTLFNPPLPDQOEISVFEHEWPSLKGNRAIALVHVNPRGGRNORAYSLPHE
AVOVALMDLAKAEAGLPLHVLGSRNRNVRKAYAGLDFHLDADPAVSLFSAHSTYS
AFKIVGHRDFDRDLRLLELTKTCVPAGSKVMIDPNEAMTSKALTRKYTRRGDL
LWVEDPILRHDDGLRTLRLHAHVTWQJNSGEYLDLOGKRLLELAHADILNVGQVTD
VMRIQTLAEALGTPISIGNTEFLEAGVIMAAVALBVEVLETSFQNFHLVEQPLEIDG
YALPADRGHGLVLESEKARGBSRRLANSELGAAEPENRLPAK"
/3870..4556
/gene="mlr7090"
/3870..4556
/gene="mlr7090"
/codon_start=1
/transl_table=11
/product="transcriptional regulator"
/protein_id="BAB53264.1"
/db_xref="GI:14026668"
/translation="MAGKATDSSPLVAARMSDINVERILEGLFDRVPAGAFISON
ELSLIVDPVAPLRDALRVLEAGLITTHPRSGIQFVRGLFETVATOFIRYVRAA
VRFPAEGDVAVNMALEIRHSRLRVERDGGKROLEBGLIGAVIYALRNP
LIDSVYRKMNYRLRLRLDRKITTPIMVRLTKLEHLDIEACGTRNAADAENKLOHFO
AAMORSLGIV"
/4684..5922
/gene="mlr7091"
/4684..5922
/gene="mlr7091"
/codon_start=1
/product="sugar ABC transporter, periplasmic binding
protein"
/protein_id="BAB53265.1"
/db_xref="GI:14026669"
/translation="MQLNKQFLMGSAALAMAGHIRGAARASATSYHHFTSSEMAS
LKITIGLQOAAADVAVTOENIPNSEYAKVSAVLAGGRDPAVMYAFEPALTYME
GLIDITLDVNGMKGRANPENRWGLSNGAIYVPAFADVMWYRAADFEFAGLAG
PRNDEFLSACRKLTIDPSKGRVAGEBAGAGGKYIDVMEAFGSITVQDPAIAG
AAAIETAITFYSDLFLEKIVVPSPA.PNSYRQIMGCFVTGTAVWVHHHTGSLNLSAM
KOGQESTFAPMPAGPKAHVAVAVAGNGIKEDNTEAMRWVSPNGQTDAAULTLEGT
GYFPASAAADDPRTGNPITYHAAVETLLEGRPLNNVGAAGSENVVNPAPQAVLTG
OSTPEQAVDRDIQGLEALR"
/5979..6932
/gene="mlr7093"
/5979..6932
/gene="mlr7093"

```



tRNA  
 /gene="tRNA-Val"  
 1105..2682  
 /product="16S rRNA"  
 gene  
 1105..2682  
 /gene="16S rRNA"  
 tRNA  
 2684..2758  
 /gene="tRNA-Leu (UUR)"  
 gene  
 2684..2758  
 /product="transfer RNA-Leu"  
 gene  
 2762..3718  
 /gene="tRNA-Leu (UUR)"  
 gene  
 2762..3718  
 /gene="NADH1"  
 CDS  
 2762..3718  
 /gene="NADH1"  
 /codon\_start=1  
 /transl\_table=2  
 /protein\_id="CAA72529.1"  
 /db\_xref="GI:2252501"  
 /db\_xref="SWISS-PROT:021325"  
 /translation="MYLNVLSLIIPIILAVAFLLERKILIGTWOLKPKPNIVGPG  
 LLOPIADAIKLFKEPLRPATSSKIMFLLPILATITLALSLMIPMPYPLVNLGV  
 LFLIAMSILAVYSILWGNASNKYALIGALRAVQTI SYEVTLAIILSLIMTNISF  
 TLSTLTTOEHMILFELPILAMMPFISTLAETNAPDITLGESELSVGFNVEYAG  
 YPRERDQMLHMLKSFPLTLALCMHVASAPAFAGVPPHM"  
 3717..3785  
 /gene="tRNA-Ile"  
 gene  
 3717..3785  
 /product="transfer RNA-Ile"  
 tRNA  
 3717..3785  
 /gene="tRNA-Ile"  
 complement(3783..3855)  
 gene  
 3856..3924  
 /product="transfer RNA-Gln"  
 tRNA  
 3856..3924  
 /gene="tRNA-Gln"  
 complement(3783..3855)  
 gene  
 3856..3924  
 /product="transfer RNA-Gln"  
 tRNA  
 3856..3924  
 /gene="tRNA-Met"  
 gene  
 3856..3924  
 /product="transfer RNA-Met"  
 gene  
 3925..4968  
 /gene="NADH2"  
 CDS  
 3925..4968  
 /gene="NADH2"  
 /codon\_start=1  
 /transl\_table=2  
 /protein\_id="CAA72530.1"  
 /db\_xref="GI:2252502"  
 /db\_xref="SWISS-PROT:021326"  
 /translation="MNPILFIIIMFTLIGTVITMISSHMLIMNGLEMFNMFPII  
 MMSHPSTEAATKYPMTQATASMLMGCVILINLYSGQMTINNSLVNYSMTTIL  
 AMKIGLAFHFVPEVTOGLTSGMILITLQKLAIPITLQIHSSLMNLNMTLAL  
 SILIGGGLNOTOLRKIMAYSSIAHMGMTIINTNLSMLNLVILMNTITMFL  
 FINSTTTTSLSLTWNSPRITTMILITLUSLGLPLSGFAPAKMIIEMTNMNL  
 LPLPTAIALMLNTFYMLITSLTSLMPTTNMKMKYKTONFIPLSPTLITLST  
 MLPLTPMMLILN"  
 4967..5034  
 /gene="tRNA-Tip"  
 gene  
 4967..5034  
 /product="transfer RNA-Tip"  
 tRNA  
 4967..5034  
 /gene="tRNA-Tip"  
 complement(5038..5106)  
 gene  
 5107..5180  
 /product="transfer RNA-Ala"  
 gene  
 5107..5180  
 /product="transfer RNA-Ala"  
 complement(5038..5106)  
 gene  
 5181..5218  
 /product="transfer RNA-Asn"  
 gene  
 5181..5218  
 /product="transfer RNA-Asn"  
 complement(5108..5180)  
 gene  
 5181..5218  
 /product="transfer RNA-Asn"  
 complement(5108..5180)  
 rep\_origin

tRNA  
 /note="origin of L-strand replication"  
 complement(5214..5274)  
 /gene="tRNA-Cys"  
 gene  
 5214..5274  
 /product="transfer RNA-Cys"  
 complement(5214..5274)  
 tRNA  
 5275..5347  
 /gene="tRNA-Cys"  
 gene  
 5275..5347  
 /product="transfer RNA-Tyr"  
 gene  
 5348..5890  
 /gene="tRNA-Tyr"  
 CDS  
 5348..5890  
 /gene="COI"  
 5349..6890  
 /gene="COI"  
 /codon\_start=1  
 /transl\_table=2  
 /protein\_id="CAA72525.1"  
 /db\_xref="GI:2252503"  
 /db\_xref="SWISS-PROT:021327"  
 /translation="MFLTRMLFSTNHRDIDGTYLLEFGAWAGVGTALSLIRALGCP  
 GTLLDDQIVNIVTAHAFIMFENYMPIDIGGFGNWLPIMLGAPDAPFRNNMSF  
 WLLPSPFLILASSWAGAGTGTVPPLACNLAAASVDITITSLDGLGISLIG  
 AINFTITILNMKPPAMTOXOTPLVWSVTVTVLILSLIPVAGITMLLDGRNLT  
 FPDAGGDDPLYOHLFWPGHPEVYIILPBGMI SHIVYSGKDEPGMGWVA  
 MSIGFLGFIYVAHMHMFTYOMDYNATVTSATMILAIPTGVKFSMLATLHGNIKW  
 SPATMIAIGFIRFLTYGGLGMLANSSLDIVLHPDTYVVAHFVHVSMAGVFAIMGC  
 FVHMFELFSGTYLNTLWAKIHFIMVGVNLFFPPOHGLGSGMPRRYSOPVAYTM  
 NYSSKGSFSLAIVMLIMFIMEAFASKREVDELPTNLMLHCCPPHYHFEF  
 AFVKV"  
 complement(6892..6960)  
 /gene="tRNA-Ser (UCN)"  
 gene  
 6961..7032  
 /product="transfer RNA-Ser"  
 complement(6892..6960)  
 tRNA  
 6961..7032  
 /gene="tRNA-Ser (UCN)"  
 complement(6892..6960)  
 gene  
 7033..7717  
 /gene="tRNA-Asp"  
 gene  
 7033..7717  
 /product="transfer RNA-Asp"  
 gene  
 7033..7717  
 /gene="tRNA-Asp"  
 complement(6892..6960)  
 CDS  
 7033..7717  
 /gene="COIT"  
 /codon\_start=1  
 /transl\_table=2  
 /protein\_id="CAA72526.1"  
 /db\_xref="GI:2252504"  
 /db\_xref="SWISS-PROT:P50687"  
 /translation="MYPDLGQFODATSPIMEELHFNHDTMLIVELISSLVYITL  
 MTKIRHSTMDAQEVEYVTLIPVILIALPRLRYMMDIENPDLITKAMGH  
 QWMSYETIEDIEDLFDYSVWPTSDIKPGELRLLEVDNLYVPMELSLRMILISSEYV  
 HSMVPSLIGKTDIIPRLNOATIMATRGLYGCGSEICSNHSPMIVLELPLAH  
 FEMDSTSNL"  
 7721..7791  
 /gene="tRNA-Lys"  
 gene  
 7721..7791  
 /product="transfer RNA-Lys"  
 gene  
 7721..7791  
 /gene="tRNA-Lys"  
 complement(7795..7995)  
 CDS  
 7796..7995  
 /gene="ATPase8"  
 7796..7995  
 /gene="ATPase8"  
 /codon\_start=1  
 /transl\_table=2  
 /protein\_id="CAA72527.1"  
 /db\_xref="GI:2252505"  
 /db\_xref="SWISS-PROT:021329"  
 /translation="MPOLDTSNWFITIVSMLSLFIIMOLKFKIKSSFSSTPCPTTMEK  
 TKHLTPWEKKTKTYLPLPSLPLP"  
 7953..8633  
 /gene="ATPase6"  
 7953..8633



	/gene="ATPase6"	gene
	/codon_start=1	gene
	/transl_table=2	tRNA
	/protein_id="CAA72528.1"	tRNA
Query Match	33.0%; Score 37.6; DB 4; Length 17056;	gene
Best Local Similarity	63.0%; Pred. No. 0.17;	gene
Matches 58; Conservative 0; Mismatches 34; Indels 0; Gaps 0;		tRNA
Oy	B accgacgcagcaagaaccgcgatctggcgtatgccatcgcccttgctgcatctc 67            Db 9090 AAGTCACCGGAAGCATCTACGAGCCCTATTACTATCATACTGCCTA 9149 	gene
Oy	68 ttaccttggccaagtcttatgatattataaa 99 	tRNA
RESULT 4	SU0238588 16507 bp DNA circular ROD 06-JUN-2000	CDS
LOCUS	SU0238588 Sciturus vulgaris complete mitochondrial genome.	gene
DEFINITION	AJ238588 AJ238588.1 GI:8347068	gene
ACCESSION	12S ribosomal RNA; 12S rRNA gene; 16S ribosomal RNA; 16S rRNA gene; ATPase 6 gene; ATPase 8 gene; ATPase subunit 6; ATPase subunit 8; COI gene; COII gene; control region; cytb gene; cytochrome b; cytochrome c oxidase subunit I; cytochrome c oxidase subunit II; cytochrome c oxidase subunit III; D-loop; mitochondrion; NADH dehydrogenase subunit 3; NADH dehydrogenase subunit 2; NADH dehydrogenase subunit 4L; NADH dehydrogenase subunit 5; NADH dehydrogenase subunit 6; NADH gene; NADH2 gene; NADH3 gene; NADH4 gene; NADH4L gene; NADH5 gene; NADH6 gene; transfer RNA-Ala; transfer RNA-Arc; transfer RNA-Asn; transfer RNA-ASP; transfer RNA-Cys; transfer RNA-Gln; transfer RNA-Gly; transfer RNA-His; transfer RNA-Ile; transfer RNA-Leu(CUU); transfer RNA-Leu(UUR); transfer RNA-Lys; transfer RNA-Met; transfer RNA-Phe; transfer RNA-Ser(AGY); transfer RNA-Tip; transfer RNA-Tyr; transfer RNA-Val; tRNA-Ala gene; tRNA-Arg gene; tRNA-Arg gene; tRNA-Asp gene; tRNA-Cys gene; tRNA-Gln gene; tRNA-Gly gene; tRNA-His gene; tRNA-Ile gene; tRNA-Leu(CUU) gene; tRNA-Leu(UUR) gene; tRNA-Lys gene; tRNA-Met gene; tRNA-Phe gene; tRNA-Ser (AGY) gene; tRNA-Trp gene; tRNA-Tyr gene; tRNA-Val gene. Eurasian red squirrel.	tRNA
SOURCE	Mitochondrion Sciturus vulgaris	gene
ORGANISM	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Rodentia; Sciurognathi; Scuridae; Sciurinae; Sciurus.	tRNA
REFERENCE	1 (bases 1 to 16507)	gene
AUTHORS	Reyes A., Glasl C., Pesole G., Catzeffls F.M. and Saccone C.	CDS
TITLE	Where do rodents fit? Evidence from the complete mitochondrial genome of Sciturus vulgaris	gene
JOURNAL	Mol. Biol. Evol. 17 (6), 979-983 (2000)	gene
MEDLINE	20293317	tRNA
REFERENCE	2 (bases 1 to 16507)	gene
AUTHORS	Reyes A.	gene
TITLE	Direct Submission	gene
JOURNAL	Submitted (27-Apr-1999) Reyes A., Consiglio Nazionale delle Ricerche (CNR), Centro Studio Mitoccondri e Metabolismo Energetico, Via Amendola 165/A, Bari, 70126, ITALY	gene
FEATURES	location/Qualifiers	gene
source	1..16507 /organism="Sciturus vulgaris" /organelle="mitochondrion" /db_xref="taxon:55149"	gene
tRNA	1..66 /gene="tRNA-Phe"	tRNA
gene	1..66 /product="transfer RNA-Phe"	gene
tRNA	67..1033 /gene="tRNA-Phe"	gene
tRNA	12S rRNA"	gene

```

tRNA      complement(5033..5102)
          /gene="trna-Ala"
          /product="transfer RNA-Ala"
          complement(5033..5102)
          /gene="trna-Ala"
          complement(5106..5177)
          /gene="trna-Asn"
          /product="transfer RNA-Asn"
          complement(5106..5177)
          /gene="trna-Asn"
          complement(5178..5209)
          /direction="left"
          complement(5210..5276)
          /gene="trna-Cys"
          /product="transfer RNA-Cys"
          complement(5210..5276)
          /gene="trna-Cys"
          complement(5277..5351)
          /gene="trna-Tyr"
          /product="transfer RNA-Tyr"
          complement(5277..5351)
          /gene="trna-Tyr"
          5352..6893
          /gene="COI"
          5352..6893
          /gene="COI"
          /codon_start=1
          /product="cytochrome c oxidase subunit I"
          /protein_id="CA893987.1"
          /db_xref="GI:8347071"
          /translation="MTINRMLSTNKHDKIGTLYLLFGANAVGTALSLILRAELGQP
          GALLGDDQLYNNVTAHAFVMEFVWPMIGFGNWLPLMTGADMAFPPMNMKS
          WLLPSPFLSLSNVAGAGTGVVYPLAGNLAAAGASVLTFTSLHACVSSILG
          AINFTITINMKPAMSOYOTPLEVSVLTAVLLSLPLAAGITMLDNRNNT
          FEDPAGGDPILYOHLEFMEFGHEVYILLPQGIISHTVYVYSGKKEPFGMGYMA
          MMSIGLFTLVANHHMTGYMDVDPRAVTSMTAATPKVKSMTATLGGNIKW
          SPAMALGFIPLFTYGGTGTGLVANSGLDYLHDHYVAHFHYLSGAVFAALMG
          PYWHPFLESGITLDNTWAKIHFTVMEGVNLTFFQHLGSGMPRRSDPDAYTGW
          NVSSGSGFISLAVIMIFMIMEAFASKREVLTVELSTNLEMLHGCPPPYHTPEEP
          IYVKA"
          complement(6897..6965)
          /gene="trna-Ser(UCN)"
          /product="transfer RNA-Ser(UCN)"
          complement(6897..6965)
          /gene="trna-Ser(UCN)"
          6969..7037
          /gene="trna-Asp"
          /product="transfer RNA-Asp"
          6969..7037
          /gene="trna-Asp"
          7038..7721
          /gene="COII"
          7038..7721
          /gene="COII"
          /codon_start=1
          /transl_table=2
          /product="cytochrome c oxidase subunit II"
          /protein_id="CA893988.1"
          /db_xref="GI:8347072"
          /translation="MAYPELGFODATSPIMEELLHFDHTLMTIVELTSSVYITIS
          LMTTLHTSTMDAQEVETITWTLLPAIILLIAPLSRLIIMDEIINDPSLTKWGH
          QWVWSEYETDEDLNFDSYMTPTSLKGEELLEVDRVVLPEMLPTRLTSSEDVL
          HSWAVPSLGLKTDALPGRLNQTATLTSTRGLLYGCGSEICGSHNSFMTVLEWPLKH
          FENWSSSML"
          7725..7792
          /gene="trna-Lys"
          /product="transfer RNA-Lys"
          7725..7792
          /gene="trna-Lys"
          7794..7997
          /gene="Atpase 8"
          7794..7997
          /gene="Atpase 8"
          CDS

```

```

Query Match
Best Local Similarity 32.5%; Score 37; DB 10; Length 16507;
Matches 58; Conservative 0; Mismatches 35; Indels 0; Gaps 0;

Oy 8 acggcagccgcaagaacgcgcatgctgcgcattccatgcgcctgcctgcatc 67
    ||| ||||| ||| | ||||| ||| | ||||| ||| | ||||| ||| |
Db 9092 AAGGCAACCGCAGCAAAATCTCAAGCACTAGCTATTAATTCGACTTGCCTTAT 9151
    ||| || ||||| ||||| ||||| ||||| ||||| ||||| |||||
Oy 68 ttaccctgtgcagccataatgatgaataa 100
    ||| || ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 9152 TTACTACTCTCCAGCATCGAATACTTAGAA 9184
    ||| || ||||| ||||| ||||| ||||| ||||| ||||| |||||

RESULT 5
AF348080
LOCUS AF348080 16968 bp DNA circular MAM 01-AUG-2001
DEFINITION Ochotona collaris mitochondrion, complete genome.
ACCESSION AF348080
VERSION AF348080.1 GI:14599763
KEYWORDS
SOURCE
ORGANISM
Ochotona collaris.
Mitochondrion Ochotona collaris
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Lagomorpha; Ochotonidae; Ochotona.
Lin.Y.-H., Waddell, P.J. and Penny, D.
Plx and vole mitochondrial genomes support both rodent monophyly
and gires
REFERENCE
1 (bases 1 to 16968)
AUTHORS
Unpublished
TITLE
2 (bases 1 to 16968)
JOURNAL
REFERENCE
1. 16968
AUTHORS
Submitted (08-FEB-2001) Institute of Molecular Biosciences, Massey
University, Palmerston North, New Zealand
LOCATION/Qualifiers
1. 16968
/organism="Ochotona collaris"
/organelle="mitochondrion"
/db_xref="taxon:134600"
1..70
/product="trna-Phe"
71..1032
/product="12S ribosomal RNA"
1033..1100
/product="trna-Val"
1101..2672
/product="16S ribosomal RNA"
2673..2748
/product="trna-Leu"
2750..3706
/codon_start=1
/transl_table=2
/product="NADH dehydrogenase subunit 1"
/protein_id="AAK71069.1"
/db_xref="GI:14599764"
/translation="MFLANTITLITIVLLAMAFITVERKILGYMOLRKPVNVGPGY
LIOPFADAVKLFYKEPLRPLTSSPFLFTAPLTALTLATMMPIPMYPPLINMGIL
LFTLATSLSLVISYILWSGASNSKVALIGALRAVADTISYEVTALITLITLIMNGS
TISLSDIDDEYKMLIIPAMPPLAMNFIISLAEINRAFPDLTGESESLVSGFVVEYAG
PFAFLFELAEYTNINMNAALTVLTGAVYDTHNPOLFTINFAIKTLLTWPLWLRAS
YPRFRDOLHMLMKNEFLPLTALCMWHSMPVFLASIPPA"
3706..3773
/product="trna-Ile"
complement(3771..3842)
/product="trna-Gln"
3843..3911
/product="trna-Met"
3912..4955
/codon_start=1
/transl_table=2
/product="NADH dehydrogenase subunit 2"

```



2000  
 1999  
 1998  
 1997  
 1996  
 1995  
 1994  
 1993  
 1992  
 1991  
 1990  
 1989  
 1988  
 1987  
 1986  
 1985  
 1984  
 1983  
 1982  
 1981  
 1980  
 1979  
 1978  
 1977  
 1976  
 1975  
 1974  
 1973  
 1972  
 1971  
 1970  
 1969  
 1968  
 1967  
 1966  
 1965  
 1964  
 1963  
 1962  
 1961  
 1960  
 1959  
 1958  
 1957  
 1956  
 1955  
 1954  
 1953  
 1952  
 1951  
 1950  
 1949  
 1948  
 1947  
 1946  
 1945  
 1944  
 1943  
 1942  
 1941  
 1940  
 1939  
 1938  
 1937  
 1936  
 1935  
 1934  
 1933  
 1932  
 1931  
 1930  
 1929  
 1928  
 1927  
 1926  
 1925  
 1924  
 1923  
 1922  
 1921  
 1920  
 1919  
 1918  
 1917  
 1916  
 1915  
 1914  
 1913  
 1912  
 1911  
 1910  
 1909  
 1908  
 1907  
 1906  
 1905  
 1904  
 1903  
 1902  
 1901  
 1900

...CCCAAGCCACAGAAATATATGAA 549

RESULT 7  
0A085910

DEFINITION			
Ovis aries cytochrome c-oxidase subunit three (COITIII) gene	1150 bp	DNA	MAM 02-OCT-1997

genes encoding mitochondrial products.  
U85910

**KEYWORDS** : sheep.

**SOURCE** :

Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
Mammalia; Eutheria

REFERENCE  
1 (bases 1 to 1130)  
Bovidae; Caprinae; Ovis.  
Bovidae

**TITLE** Nucleotide sequence of ovine mitochondrial genes for tRNA-Gly and subunit three of cytochrome c oxidase

REFERENCE  
2 (bases 1 to 1130)  
AUTHORS  
Henderson C

Submitted (21-JAN-1997) Anim. Breed. Genet.

FEATURES	1990/1991	Germany
source	location/Qualifiers	
	1. 1130	

```

/organelle="mitochondrion"
/strain="modified"

```

```

/uv_xref="taxon:9940"
/tissue_type="liver"

```

	/gene="COLII"
CDS	75..878

```

/codon_start=1
/transl_table=3

```

```

/product_cytochrome_c-oxidase subunit three"
/protein_id="AAB71760.1"

```

```
/translation="MTHQTHAYHMVNPSPWPLGALSALMLTSGLMFHFNSTALL  
LGLTTNMLTWQWRDVIRESFGHHTPAVKCIKVCMTITTCVETDDE
```

QALFTIALGVYFTLLQASEYYEAPFTISDGVGSGTTFVATGFHGLHVITIGSFEFLIV  
FFROLKEHETSSUHECTEEA

```
069 1027  
/gene="tRNA-Gly gene"  
050 027
```

BASE COUNT	307	a	324	c	165	g	334	t
GCCTT	/gene="tRNA-Gly gene"							

Best Local Similarity	60.28;	DB 4;	Length 1130;
Matches	59; Conservative	Pred. No. 0.59;	

2 tgcgtccacggcgacgcgaagaaacgcgcgatctgcctacgaattgccatcanncttgcct 61

520 TCATGAGGGAACCGTTACCAATGTTACAAACCCATTTCATTTACCAATCGACCTAGGCG 585

586 TGTACTTTACACATTGTTACAGCCCAACCAAGCAATTTATTTT

8 SMIT

Accession	Size	Library	Year
AF010406	16616 bp	DNA circular	MAM 13-DEC 2000
U05			
U0400			

1

DEFINITION Ovis aries complete mitochondrial genome.  
 ACCESSION AF010406  
 VERSION AF010406.1 GI:3445513  
 KEYWORDS  
 SOURCE sheep.  
 ORGANISM Mitochondrion Ovis aries  
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 Mammalia; Eutheria; Cetartiodactyla; Ruminantia; Pecora; Bovidea;  
 Bovidae; Caprinae; Ovis.  
 REFERENCE  
 AUTHORS 1 (bases 1 to 16616)  
 TITLE Hiendleder, S., Mainz, K., Planke, Y. and Lewalski, H.  
 Analysis of mitochondrial DNA indicates that domestic sheep are  
 derived from two different ancestral maternal sources: no evidence  
 for contributions from urial and argali sheep  
 J. Hered. 89 (2), 113-120 (1998)  
 JOURNAL 2 (bases 1 to 16616)  
 MEDLINE 98203262  
 REFERENCE  
 AUTHORS A low rate of replacement substitutions in two major Ovis aries  
 TITLE Hiendleder, S.  
 Mitochondrial genomes  
 Anim. Genet. 23 (2), 116-122 (1998)  
 JOURNAL 3 (bases 1 to 16616)  
 MEDLINE 98364427  
 REFERENCE  
 AUTHORS Hiendleder, S., Lewalski, H., Wassmuth, R. and Janke, A.  
 TITLE The complete mitochondrial DNA sequence of the domestic sheep (Ovis  
 aries) and comparison with the other major ovine haplotype  
 J. Mol. Evol. 47 (4), 441-448 (1998)  
 JOURNAL 5 (bases 1 to 16616)  
 MEDLINE 98440761  
 REFERENCE  
 PUBMED 9767689  
 AUTHORS 4 (bases 1 to 16616)  
 TITLE Hiendleder, S., Wassmuth, R. and Lewalski, H.  
 Direct Submission  
 Submitted (26-JUN-1997) Animal Breeding and Genetics,  
 Justus-Liebig-University, Ludwigstr. 21B, Giessen 35390, Germany  
 REFERENCE  
 AUTHORS 5 (bases 1 to 16616)  
 TITLE Hiendleder, S., Wassmuth, R. and Lewalski, H.  
 Direct Submission  
 Submitted (19-AUG-1998) Animal Breeding and Genetics,  
 Justus-Liebig-University, Ludwigstr. 21B, Giessen 35390, Germany  
 REMARK  
 COMMENT On Aug 21, 1998 this sequence version replaced GI:3366622.  
 FEATURES  
 source  
 1..16616  
 /organism="Ovis aries"  
 /organelle="mitochondrion"  
 /strain="Merinolandschaf"  
 /db\_xref="taxon:9940"  
 /tissue\_type="liver"  
 /note="haplotype B (European type)"  
 1..68  
 /product="tRNA-Phe"  
 69..1026  
 /product="12S ribosomal RNA"  
 1027..1093  
 /product="tRNA-Val"  
 1094..2667  
 /product="tRNA-16S ribosomal RNA"  
 2668..2742  
 /note="tRNA"  
 /product="tRNA-Leu"  
 2745..3699  
 /gene="ND1"  
 2745..3699  
 /gene="ND1"  
 /note="ND1"  
 /note="tRNA stop codon is completed by the addition of 3' A  
 residues to the mRNA"  
 /transl\_except=(pos:3699,aa:TERM)  
 /transl\_table=2  
 /product="NADH dehydrogenase subunit 1"  
 /protein\_id="GI:3366623"  
 /db\_xref="GI:3366623"  
 /translation="MFMINVLILITLILAAVAFLITLVERKVLGYMOPKRGPNVAVPYG

LLPITADIKITKEPLRPATSSISMFILAPLITLALTMNIPLEMPYPLINMLGV  
 LFLMLASSLAVYSITLWSCGANSKVALIGALRAVAOTISYEVALIILISVLNMGSE  
 TLSTLITTOEQVWLTLPAPMLAMNPELSTLAETNAPPDITGESEVSGFVVEYAAQ  
 PFALFEAAEYANILIMNIFTTLLPLGAFNHPYPELYTINFTIKSLLSITPLMIRAS  
 YPRFRYDOLMLMKNFPLPLALCMWHVSPILLSIPPQT  
 3701..3769  
 /product="tRNA-Ile"  
 complement(3767..3838)  
 /product="tRNA-Gln"  
 3841..3909  
 /product="tRNA-Met"  
 3910..4951  
 /gene="ND2"  
 3910..4951  
 /gene="ND2"  
 /note="tRNA stop codon is completed by the addition of 3' A  
 residues to the mRNA"  
 /codon\_start=1  
 /transl\_except=(pos:4951,aa:TERM)  
 /transl\_table=2  
 /product="NADH dehydrogenase subunit 2"  
 /protein\_id="GI:3366624"  
 /db\_xref="GI:3366624"  
 /translation="MAPILILITMTVAGITIVISRWLLMIGFEMNLAITPIM  
 MKKINRATVASTKFTFDTOSMLMAIINIMFGQWYMLKLPNMAEMLTAL  
 AKLGMAPPHFWPEVTOGIPUSGLILITWOKLAPSVLQIILPINDLILITSL  
 SITIGMGGLNOTOLKIMAYSSIAHMGWMAVLYNPTMLNLIITIMTSMPTL  
 FMANSYTTTSLSHTNKAPIMTIVLITLISMGILPISGMPKRMKTIOMTKNDST  
 ILPTLMAITALLNLYFYMRVLTALTEPSTNNKMKWQSPPTTKRMTLTPTMTVLS  
 MLPLPILISILE"  
 4952..5018  
 /product="tRNA-Trp"  
 complement(5020..5088)  
 /product="tRNA-Ala"  
 complement(5090..5162)  
 /product="tRNA-Asn"  
 5163..5194  
 /note="for L-strand"  
 complement(5195..5262)  
 /product="tRNA-Cys"  
 complement(5263..5330)  
 /product="tRNA-Tyr"  
 5332..6876  
 /gene="COI"  
 5332..6876  
 /gene="COI"  
 /codon\_start=1  
 /transl\_table=2  
 /product="cytochrome c-oxidase subunit 1"  
 /protein\_id="GI:3366625"  
 /db\_xref="GI:3366625"  
 /translation="MFIRKRLFSRNHMDIGTIVLLFGAGAGVGTALSLIRALGQP  
 GTLLDDQIVYIVTAHAFVMTFPMWPMIMGCGMVLVPLMGAPMAPPNNNSF  
 MLPPSFLILLASSWENAGAGTGTVPPLAGNLAHAGASVDITLSILHAGVSLIG  
 AINFTYTLINNKPRAMSQYQPLFVWSVLITVALLSLDFVLAAGITIMLTDNNLMT  
 PEDPAGGDPILYCHLFWFPGHPEVYILILFEGEMISHIVYISGKEPFGYGMVA  
 MMSIGELGIYWAHMFVMDVDVDFRAFTSATMIIAIPGVKVFSLATLHGNKTKM  
 SPAMMAALGFIETLVGGTIGVIVANSLLDVLDTYVVAHRYVLSKAVPAIKMG  
 FVHMFPLPSGTYLNDPTNPAKIHFAIMFVGNMTFFPQHLGSGMRRYSQYPPAYTMM  
 NTISSGFSISLAVMLMTIITWEAFASKREVLYVDLTTNLEMLNCCPPHYHTFEP  
 TVYNLX"  
 complement(6874..6944)  
 /note="UCN"  
 /product="tRNA-Ser"  
 6950..7017  
 /product="tRNA-Asp"  
 7019..7702  
 /gene="COI1"  
 7019..7702  
 /gene="COI1"  
 /codon\_start=1  
 /transl\_table=2  
 /product="cytochrome c-oxidase subunit 2"











[illegible]





Query Match	30.5%;	Score 34.8;	DB 4;	Length 783;
Best Local Similarity	57.3%;	Pred. No. 1;		
Matches 63;	Conservative 0;	Mismatches 47;	Indels 0;	Gaps 0;

QY	Db	QY
2	tcgtccagcgagccgcgaagaacccgcatgtgactagcgatgcatcgcccttgct	61
111		
452	tcattaaagaaatgacacccacatacttcaagccctattcatattacatgcgacctaagg	511
111		
62	gaatccttaacccgtgtgcgaagcctatgatatactgaatcgcgtccaaacc	111
111		
512	ttttacttcacatctctccaagcctctagaaattttgcgaagccctttatc	561
111		

```
Search completed: January 10, 2002, 00:56:35
Job time: 3083 sec
```

Db 452 TCATAGAGGAATATGCCAACCATCTACTTCAAGGCGCTATTTCATTACCATCGACATAGGGG 511  
 QY 62 ggatccttaccctgtgcgaagcgtatgatatatgaatcgccatacc 111  
 Db 512 TATACCTTCACACTACTCCCAAGCTCCTCAAGATATATTCGAGAGACCGCTTTTAC 561

LOCUS	783 bp	DNA	MAM	26-JUL-1999
DEFINITION	<i>Gazella thomsonii</i> cytochrome oxidase subunit III (cox iii) gene			
ACCESSION	mitochondrial gene encoding mitochondrial protein, partial cds.			
VERSION	AF030473			
FEATURES	AF030473.1	GI:2731923		

REFERENCE	AUTHORS	TITLE
1 (bases 1 to 783)	Rebholz, W. and Harley, E.	Phylogenetic relationships in the bovid subfamily Bovinae

JOURNAL OF CLIMATE AND DOCKLAND ANTHROPOLOGY  
MOL. PHYLOGENET. EVOL. 12 (2), 87-94 (1999)  
99310773

PUBMED	10381312
REFERENCE	2 (bases 1 to 783)
AUTHORS	Rehbolz/W.E.R. and Harley,E.H.
TITLE	Direct Submission
JOURNAL	Submitted (21-Oct-1997) Chemical Pathology, University of Cape Town Medical School, Observatory, Cape Town 7925, South Africa
FEATURES	Location/Qualifiers

```

gene
CDS
    /map=89/0-9/52 relative to the cow (Bos taurus) sequence
    /note="From blood sample"
    1..>783
    /gene="cox 11"
    1..>783

```

BASE COUNT	ORIGIN
217 a	1
238 c	2
107 g	3
221 t	4

Query Match	30.5%	Score 34.8;	DB 4;	Length 783;
Best Local Similarity	57.3%;	Pred. No. 1;		
Matches 63;	Conservative 0;	Mismatches 47;	Indels 0;	Gaps 0;

OY    2 tcgtcacgagcgacgcgaagaaccggatgctagcattgccatgcagccttgct 61



Thu Jan 10 08:15:40 2002

us-09-712-768-5.rai

Page 2

```

; Patent No. 5840695
; GENERAL INFORMATION:
; APPLICANT: FRANK, GLENN R.
; APPLICANT: HUNTER, SHIRLEY WU
; APPLICANT: WALLENFELS, LYNDA
; TITLE OF INVENTION: NOVEL ECTOPARASITE SALIVA PROTEINS
; NUMBER OF SEQUENCES: 107
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Sheridan Ross P.C.
; STREET: 1700 Lincoln Street, Suite 3500
; CITY: Denver
; STATE: Colorado
; COUNTRY: U.S.A.
; ZIP: 80203
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patentin Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US-08/630,822A
; FILING DATE: 11-APR-1996
; CLASSIFICATION: 435
; ATTORNEY/AGENT INFORMATION:
; NAME: CONNELL, GARY J.
; REGISTRATION NUMBER: 32,020
; REFERENCE/DOCKET NUMBER: 2618-17-C3
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (303) 863-9700
; TELEFAX: (303) 863-0223
; INFORMATION FOR SEQ ID NO: 60:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 157 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: protein
; US-08-630-822A-60

alignment_scores:
  Quality: 53.00      Length: 27
  Ratio: 2.944      Gaps: 1
  Percent Similarity: 66.667      Percent Identity: 40.741

alignment_block:
  US-09-712-768-5/rev x US-08-630-822A-60
  Align seg 1/1 to: US-08-630-822A-60 from: 1 to: 157

80 TGGCAGCGGTAAAGATCCAGCCGCGATGCGATCGCTAGCCCAAT 31
   ||| : : : : : ||| : : : : : ||| : : : : : |||
33 TrpAspLysLeuProPheValProArgProSerLeuValAlaAspPro1 49
30 CGCGGTTTCTTTCGCGTCCGCGCGTGAGACA 2
   | : : : : | | | | | : : : : | | |
49 ethAlaPhe.CysLysArgLysProArg 58

seq_name: /cgn2_6/ptodata/2/laa/5B_COMB.pep:US-09-005-069-60
seq_documentation_block:
; Sequence 60, Application US/09005069
; Patent No. 5932470
; GENERAL INFORMATION:
; APPLICANT: FRANK, GLENN R.
; APPLICANT: HUNTER, SHIRLEY WU
; APPLICANT: WALLENFELS, LYNDA
; TITLE OF INVENTION: NOVEL ECTOPARASITE SALIVA PROTEINS
; NUMBER OF SEQUENCES: 107
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Sheridan Ross P.C.
; STREET: 1700 Lincoln Street, Suite 3500
```

```

; CITY: Denver
; STATE: Colorado
; COUNTRY: U.S.A.
; ZIP: 80203
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patentin Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/09/005,069
; FILING DATE:
; CLASSIFICATION:
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 08/630,822
; FILING DATE: 11-APR-1996
; ATTORNEY/AGENT INFORMATION:
; NAME: CONNELL, GARY J.
; REGISTRATION NUMBER: 32,020
; REFERENCE/DOCKET NUMBER: 2618-17-C3
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (303) 863-9700
; TELEFAX: (303) 863-0223
; INFORMATION FOR SEQ ID NO: 60:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 157 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: protein
; US-09-005-069-60

alignment_scores:
  Quality: 53.00      Length: 27
  Ratio: 2.944      Gaps: 1
  Percent Similarity: 66.667      Percent Identity: 40.741

alignment_block:
  US-09-712-768-5/rev x US-09-005-069-60
  Align seg 1/1 to: US-09-005-069-60 from: 1 to: 157

80 TGGCAGCGGTAAAGATCCAGCCGCGATGCGATCGCTAGCCCAAT 31
   ||| : : : : : ||| : : : : : ||| : : : : : |||
33 TrpAspLysLeuProPheValProArgProSerLeuValAlaAspPro1 49
30 CGCGGTTTCTTTCGCGTCCGCGCGTGAGACA 2
   | : : : : | | | | | : : : : | | |
49 ethAlaPhe.CysLysArgLysProArg 58

seq_name: /cgn2_6/ptodata/2/laa/5B_COMB.pep:US-09-029-213B-22
seq_documentation_block:
; Sequence 22, Application US/09029213B
; Patent No. 6180098
; GENERAL INFORMATION:
; APPLICANT: CHRISTIAN, Peter D.
; TITLE OF INVENTION: RECOMBINANT HELICOVERA BACULOVIRUSES
; NUMBER OF SEQUENCES: 27
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: McDermott, Will & Emery
; STREET: 600 13th Street, NW
; CITY: Washington
; STATE: District of Columbia
; COUNTRY: USA
; ZIP: 20005
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patentin Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
```

Thu Jan 10 08:15:40 2002

us-09-712-768-5.rai

APPLICATION NUMBER: US/09/029,213B  
FILING DATE: 31-AUG-1998  
CLASSIFICATION: 424  
ATTORNEY/AGENT INFORMATION:  
NAME: Joseph Hyosuk Kim  
REGISTRATION NUMBER: 41,425  
REFERENCE/DOCKET NUMBER: 50179-048  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: 202-756-8000  
TELEFAX: 202-756-8087  
INFORMATION FOR SEQ ID NO: 22:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 181 amino acids  
TYPE: amino acid  
TOPOLOGY: linear  
MOLECULE TYPE: protein  
US-09-029-213B-22

alignment\_scores:  
Quality: 51.00 Length: 36  
Ratio: 2.125 Gaps: 0  
Percent Similarity: 66.667 Percent Identity: 36.111

alignment\_block:

US-09-712-768-5/rev x US-09-029-213B-22 ..

Align seg 1/1 to: US-09-029-213B-22 from: 1 to: 181

113 TCGATGACGACGATTCATATATTCATAGCGCTGGCAGGGAAGAT 64  
||||| :|||:||||| :|||:|||||:|||||:|||||:|||||  
6 SerLeuLeuThrLeuSerLeuLeuGlnValValItyrHisAsnValArgI 22  
63 CGAGCCAGCGCGATGCGCATCGCTAGCGCATGCGGCTTTCTGGCGT 14  
|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|  
22 eGIuYsGIuTYrMeTAlAllePheGluHisIleArgIleAlaLeuPro 39  
13 CGCGCTGG 6  
|| |||  
39 eGIuYrTP 41

seq\_name: /cgn2\_6/ptodata/2/1aa/5A\_COMB.pep:US-08-026-138E-3

seq\_documentation\_block:

; Sequence 3, Application US/08026138E  
; Patent No. 5502166

; GENERAL INFORMATION:

; APPLICANT: Masayoshi MISHINA

; TITLE OF INVENTION: NOVEL PROTEINS AND GENES CODING THE SAME

; NUMBER OF SEQUENCES: 19

; CORRESPONDENCE ADDRESS:

; ADDRESSEE: Nishiohata Residence 1-107

; STREET: 5214, Nishiohata-machi

; CITY: Niigata-shi

; STATE: Niigata-ken

; COUNTRY: JAPAN

; ZIP: 951

; COMPUTER READABLE FORM:

; MEDIUM TYPE: Diskette, 3.50 inch, 1.44 MB storage

; COMPUTER: IBM Compatible

; OPERATING SYSTEM: MS-DOS V.5

; SOFTWARE: Word Perfect 5.1

; CURRENT APPLICATION DATA:

; APPLICATION NUMBER: US/08/026,138E

; FILING DATE: 26-FEB-1993

; PRIOR APPLICATION DATA:

; APPLICATION NUMBER: JP 39563/1992

; FILING DATE: 26-FEB-1992

; APPLICATION NUMBER: JP 173155/1992

; FILING DATE: 30-JUN-1992

; APPLICATION NUMBER: JP 215017/1992

; FILING DATE: 12-AUG-1992

; APPLICATION NUMBER: JP 303878/1992

FILING DATE: 13-NOV-1992  
ATTORNEY/AGENT INFORMATION:  
NAME: Hamburg, C. Bruce  
REGISTRATION NUMBER: 22,389  
REFERENCE/DOCKET NUMBER: F-4551  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: (212) 986-2340  
TELEFAX: (212) 953-7733  
INFORMATION FOR SEQ ID NO: 3:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 1239 amino acids  
TYPE: amino acid  
STRANDEDNESS: single strand  
TOPOLOGY: linear  
MOLECULE TYPE: protein  
ORIGINAL SOURCE:  
ORGANISM: mouse  
TISSUE TYPE: brain  
PUBLICATION INFORMATION:  
AUTHORS: Masayoshi MISHINA  
TITLE: NOVEL PROTEINS AND GENES CODING THE SAME  
RELEVANT RESIDUES IN SEQ ID NO: 3: FROM 1 to 1239  
US-08-026-138E-3

alignment\_scores:  
Quality: 50.50 Length: 43  
Ratio: 2.806 Gaps: 2  
Percent Similarity: 41.860 Percent Identity: 32.558

alignment\_block:

US-09-712-768-5 x US-08-026-138E-3 ..

Align seg 1/1 to: US-08-026-138E-3 from: 1 to: 1239

3 CGTCCAGCGCGACCGCAAGAAACCGCGATGGCCT ..... 38  
||| ||||| |||:|:| |||  
1007 ArgThrSerArgProGluArgAlaLeuProGluArgSerLeu 1023  
39 ....AGCGATGGCAT.....C 51  
|||:|:| |||||  
1023 uHISAlAlHISCySHISYrSerSerPheProAlaGluArgSerGlyA 1040  
52 GGCGCTGCGTGCATCTTACCGCTGCGCA 80  
|||||:|:| |||:|:| |||  
1040 TgProPheLeuProLeuPheProGluPro 1049

seq\_name: /cgn2\_6/ptodata/2/1aa/6B\_COMB.pep:US-09-159-106-13

seq\_documentation\_block:

; Sequence 13, Application US/09159106  
; Patent No. 6284509

; GENERAL INFORMATION:

; APPLICANT: Ferrer, Pau

; APPLICANT: Halkier, Torben

; APPLICANT: Hedegaard, Lisbeth

; TITLE OF INVENTION: An enzyme with -1,3-Glucanase

; FILE REFERENCE: 4693,204-US

; CURRENT APPLICATION NUMBER: US/09/159,106

; EARLIER FILING DATE: 1998-09-23

; EARLIER FILING DATE: 1996-12-04

; EARLIER FILING DATE: 1996-08-23

; EARLIER FILING DATE: 1997-04-14

; EARLIER FILING DATE: 1997-04-14

; NUMBER OF SEQ ID NOS: 15

; SOFTWARE: FastSeq for windows Version 3.0

; SEQ ID NO 13

; LENGTH: 303

; TYPE: PRT





Thu Jan 10 08:15:40 2002

us-09-712-768-5.ra1

Page 5

```

APPLICANT: California Institute of Technology
TITLE OF INVENTION: PLANTS HAVING MODIFIED RESPONSE TO ETHYLENE
NUMBER OF SEQUENCES: 50
CORRESPONDENCE ADDRESS:
  ADDRESSEE: Richard F. Trecartin
  STREET: 3400 Embarcadero Center, Suite 3400
  CITY: San Francisco
  STATE: California
  COUNTRY: USA
  ZIP: 94111
COMPUTER READABLE FORM:
  MEDIUM TYPE: floppy disk
  OPERATING SYSTEM: PC-DOS/MS-DOS
  SOFTWARE: Patentin Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
  APPLICATION NUMBER: US/08/484,101B
  FILING DATE: 07-JUN-1995
  CLASSIFICATION: 800
  PRIOR APPLICATION DATA:
    APPLICATION NUMBER: PCT/US94/
    FILING DATE: 01-JUL-1994
    CLASSIFICATION: 800
    PRIOR APPLICATION NUMBER: US 08/086,555
    FILING DATE: 01-JUL-1993
  ATTORNEY/AGENT INFORMATION:
    NAME: Trecartin, Richard F.
    REGISTRATION NUMBER: 31,801
    REFERENCE/DOCKET NUMBER: A-57515-2/RFT
    TELECOMMUNICATION INFORMATION:
      TELEPHONE: (415) 781-1989
      TELEFAX: (415) 398-3249
  INFORMATION FOR SEQ ID NO: 36:
    SEQUENCE CHARACTERISTICS:
      LENGTH: 635 amino acids
      TYPE: amino acid
      TOPOLOGY: linear
      MOLECULE TYPE: protein
      US-08-484-101B-36

alignment_scores:
  Quality: 49.00      Length: 28
  Ratio: 2.579       Gaps: 2
  Percent Similarity: 67.857   Percent Identity: 39.286

alignment_block:
  US-09-712-768-5 x US-08-484-101B-36 ..
Align seg 1/1 to: US-08-484-101B-36 from: 1 to: 635
  2 TCGTCACGCGCAGCCGACGAGAAACCGGATGGCTAGCGATTGCCATC 51
  ||| |||:||||| |||:||||| |||:|||||
  480 SerIleGluAlaSerValAlaIysPro.GluTyrAlaArgAspCysHisP 496
  52 GGCCTTGCGTGCATCTTACCTGTGCCAAGC 83
  ||| |||:||||| |||:||||| |||:|||||
  496 roPro.....GluMetPheProMetProSer 504

seq_name: /cgn2_6/ptodata/2/1aa/5b_COMB.pep:US-08-484-101B-50
seq_documentation_block:
  Sequence 50, Application US/08484101B
  Patent No. 5824868
  GENERAL INFORMATION:
  APPLICANT: California Institute of Technology
  TITLE OF INVENTION: PLANTS HAVING MODIFIED RESPONSE TO ETHYLENE
  NUMBER OF SEQUENCES: 50
  CORRESPONDENCE ADDRESS:

```

```

  ADDRESSEE: Richard F. Trecartin
  STREET: 3400 Embarcadero Center, Suite 3400
  CITY: San Francisco
  STATE: California
  COUNTRY: USA
  ZIP: 94111
COMPUTER READABLE FORM:
  MEDIUM TYPE: floppy disk
  OPERATING SYSTEM: PC-DOS/MS-DOS
  SOFTWARE: Patentin Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
  APPLICATION NUMBER: US/08/484,101B
  FILING DATE: 07-JUN-1995
  CLASSIFICATION: 800
  PRIOR APPLICATION DATA:
    APPLICATION NUMBER: PCT/US94/
    FILING DATE: 01-JUL-1994
    CLASSIFICATION: 800
    PRIOR APPLICATION NUMBER: US 08/086,555
    FILING DATE: 01-JUL-1993
  ATTORNEY/AGENT INFORMATION:
    NAME: Trecartin, Richard F.
    REGISTRATION NUMBER: 31,801
    REFERENCE/DOCKET NUMBER: A-57515-2/RFT
    TELECOMMUNICATION INFORMATION:
      TELEPHONE: (415) 781-1989
      TELEFAX: (415) 398-3249
  INFORMATION FOR SEQ ID NO: 50:
    SEQUENCE CHARACTERISTICS:
      LENGTH: 635 amino acids
      TYPE: amino acid
      TOPOLOGY: linear
      MOLECULE TYPE: protein
      US-08-484-101B-50

alignment_scores:
  Quality: 49.00      Length: 28
  Ratio: 2.579       Gaps: 2
  Percent Similarity: 67.857   Percent Identity: 39.286

alignment_block:
  US-09-712-768-5 x US-08-484-101B-50 ..
Align seg 1/1 to: US-08-484-101B-50 from: 1 to: 635
  2 TCGTCACGCGCAGCCGACGAGAAACCGGATGGCTAGCGATTGCCATC 51
  ||| |||:||||| |||:||||| |||:|||||
  480 SerIleGluAlaSerValAlaIysPro.GluTyrAlaArgAspCysHisP 496
  52 GGCCTTGCGTGCATCTTACCTGTGCCAAGC 83
  ||| |||:||||| |||:||||| |||:|||||
  496 roPro.....GluMetPheProMetProSer 504

seq_name: /cgn2_6/ptodata/2/1aa/5b_COMB.pep:US-08-714-524D-36
seq_documentation_block:
  Sequence 36, Application US/08714524D
  Patent No. 6294716
  GENERAL INFORMATION:
  APPLICANT: Meyerowitz, Elliott M
  APPLICANT: Blecker, Anthony B
  TITLE OF INVENTION: PLANTS HAVING MODIFIED RESPONSE TO ETHYLENE
  FILE REFERENCE: a-57515-4
  CURRENT APPLICATION NUMBER: US/08/714,524D
  CURRENT FILING DATE: 1996-09-16
  NUMBER OF SEQ ID NOS: 56
  SOFTWARE: Patentin Ver. 2.1
  SEQ ID NO 36

```

Thu Jan 10 08:15:40 2002

us-09-712-768-5.ra1

```

; LENGTH: 635
; TYPE: PRT
; ORGANISM: Lycopersicon esculentum
US-08-714-524D-36

alignment_scores:
  Quality: 49.00      Length: 28
  Ratio: 2.579        Gaps: 2
  Percent Similarity: 67.857  Percent Identity: 39.286
alignment_block:
US-09-712-768-5 x US-08-714-524D-36 ..
Align seg 1/1 to: US-08-714-524D-36 from: 1 to: 635
2 TCCTCCACGGCGACCGCAAGAAACCGGATGCGCTAGCGATTCGCATC 51
|||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||
480 SerIleGluAlaSerValAlaLysPro.GluTyrAlaIaArgAspCysHisP 496
52 GGCCTTGCTGGATCTTACCCCTGTCGCAAGC 83
|||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||
496 roPro.....GluMetPheProMetProSer 504

seq_name: /cgn2_6/ptodata/2/1aa/6B-COMB.pep:US-08-714-524D-50

seq_documentation_block:
; Sequence 50, Application US/08714524D
; Patent No. 6294716
; GENERAL INFORMATION:
; APPLICANT: Meyerowitz, Elliott M
; APPLICANT: Chang, Caren
; TITLE OF INVENTION: PLANTS HAVING MODIFIED RESPONSE TO ETHYLENE
; FILE REFERENCE: a-57515-4
; CURRENT APPLICATION NUMBER: US/08/714,524D
; CURRENT FILING DATE: 1996-09-16
; NUMBER OF SEQ ID NOS: 56
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 50
; LENGTH: 635
; TYPE: PRT
; ORGANISM: Lycopersicon esculentum
US-08-714-524D-50

alignment_scores:
  Quality: 49.00      Length: 28
  Ratio: 2.579        Gaps: 2
  Percent Similarity: 67.857  Percent Identity: 39.286
alignment_block:
US-09-712-768-5 x US-08-714-524D-50 ..
Align seg 1/1 to: US-08-714-524D-50 from: 1 to: 635
2 TCCTCCACGGCGACCGCAAGAAACCGGATGCGCTAGCGATTCGCATC 51
|||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||
480 SerIleGluAlaSerValAlaLysPro.GluTyrAlaIaArgAspCysHisP 496
52 GGCCTTGCTGGATCTTACCCCTGTCGCAAGC 83
|||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||
496 roPro.....GluMetPheProMetProSer 504

seq_name: /cgn2_6/ptodata/2/1aa/6B-COMB.pep:US-08-460-269C-2

seq_documentation_block:
; Sequence 2, Application US/08460269C
; Patent No. 6197548
; GENERAL INFORMATION:
; APPLICANT: CLARE, JEFFREY J.
; APPLICANT: ROMANOS, MICHAEL A.
; TITLE OF INVENTION: EXPRESSION OF HETEROLOGOUS PROTEIN IN
```

```

; NUMBER OF SEQUENCES: 17
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Millen, White, Zelano & Branigan, P.C.
; STREET: 2200 Clarendon Blvd., Suite 1400
; CITY: ARLINGTON
; STATE: VA
; COUNTRY: USA
; ZIP: 22201
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/460,269C
; FILING DATE: 02-Jun-1995
; ATTORNEY/AGENT INFORMATION:
; NAME: Lebowitz, Richard M.
; REGISTRATION NUMBER: 37,067
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (703) 243-6410
; TELEFAX: (703) 243-6410
; INFORMATION FOR SEQ ID NO: 2:
; LENGTH: 910 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: protein
; SEQUENCE DESCRIPTION: SEQ ID NO: 2:
US-08-460-269C-2

alignment_scores:
  Quality: 49.00      Length: 30
  Ratio: 2.450        Gaps: 1
  Percent Similarity: 66.667  Percent Identity: 36.667
alignment_block:
US-09-712-768-5 x US-08-460-269C-2 ..
Align seg 1/1 to: US-08-460-269C-2 from: 1 to: 910
4 GTCCACGGCGACCGCAAGAAACCGGATGCGCTAGCGATTCGCATCG 53
|||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||
863 LeuArgGlyThrArgAlaGluLeuGlyLeuGlyMetAlaAlaLeuGI 879
54 CCTGGCTGGATCTTACCCCTGTCGCAAGCCTTGAAATAT 93
|||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||
879 yArgGly.....HisSerLeuTyrAlaSerTyrGluTyr 890

seq_name: /cgn2_6/ptodata/2/1aa/6B-COMB.pep:US-08-460-269C-4

seq_documentation_block:
; Sequence 4, Application US/08460269C
; Patent No. 6197548
; GENERAL INFORMATION:
; APPLICANT: CLARE, JEFFREY J.
; APPLICANT: ROMANOS, MICHAEL A.
; TITLE OF INVENTION: EXPRESSION OF HETEROLOGOUS PROTEIN IN
; MEDIUM TYPE: YEAST
; NUMBER OF SEQUENCES: 17
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Millen, White, Zelano & Branigan, P.C.
; STREET: 2200 Clarendon Blvd., Suite 1400
; CITY: ARLINGTON
; STATE: VA
; COUNTRY: USA
; ZIP: 22201
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
```

```
;
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patentin Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/460,269C
; FILING DATE: 02-Jun-1995
; ATTORNEY/AGENT INFORMATION:
; NAME: Lebovitz, Richard M.
; REGISTRATION NUMBER: 37,067
; REFERENCE/DOCKET NUMBER: Popov-2
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (703) 243-6333
; TELEFAX: (703) 243-6410
; INFORMATION FOR SEQ ID NO: 4:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 911 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: protein
; SEQUENCE DESCRIPTION: SEQ ID NO: 4:
US-08-460-269C-4

alignment_scores:
      Quality: 49.00      Length: 30
      Ratio: 2.450      Gaps: 1
Percent Similarity: 66.667      Percent Identity: 36.667

alignment_block:
US-09-712-768-5 x US-08-460-269C-4  ..

Align seg 1/1  to: US-08-460-269C-4  from: 1  to: 911

      4  GTCCACGGCGACCGCAAGAAACCGCATTTGGCTAGCGATGCCATCGG  53
          :::::| | | | | :::::| | | | | | | | | | | | | | |
      864  LeuArgGlyThrArgAlaGluLeuGlyLeuGlyMetAlaAlaLeuGI  880
          | | | | | :::::| | | | | | | | | | | | | | |
      54  CCTTGGCTGATCTTTACCTGTGCGAAGCTATGAATAT  93
          | | | | | :::::| | | | | | | | | | | | | | |
      880  YArgGly.....HisSerLeuTyrAlaSerTyrGluTyr  891

seq_name: /cgn2.6/prodata/2/laa/5B_COMB.pep:US-08-460-269C-6

seq_documentation_block:
; Sequence 6, Application US/08460269C
; Patent No. 6197548
; GENERAL INFORMATION:
; APPLICANT: CLARE, JEFFREY J.
; ROMANOS, MICHAEL A.
; TITLE OF INVENTION: EXPRESSION OF HETEROLOGOUS PROTEIN IN
; YEAST
; NUMBER OF SEQUENCES: 17
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Millen, White, Zelano & Branigan, P.C.
; STREET: 2200 Clarendon Blvd., Suite 1400
; CITY: ARLINGTON
; STATE: VA
; COUNTRY: USA
; ZIP: 22201
; COMPUTER READABLE FORM:
; MEDIUM TYPE: floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patentin Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/460,269C
; FILING DATE: 02-Jun-1995
; ATTORNEY/AGENT INFORMATION:
; NAME: Lebovitz, Richard M.
; REGISTRATION NUMBER: 37,067
; REFERENCE/DOCKET NUMBER: Popov-2
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (703) 243-6333
; TELEFAX: (703) 243-6410
;
```

```
;
; INFORMATION FOR SEQ ID NO: 6:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 922 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: protein
; SEQUENCE DESCRIPTION: SEQ ID NO: 6:
US-08-460-269C-6

alignment_scores:
      Quality: 49.00      Length: 30
      Ratio: 2.450      Gaps: 1
Percent Similarity: 66.667      Percent Identity: 36.667

alignment_block:
US-09-712-768-5 x US-08-460-269C-6  ..

Align seg 1/1  to: US-08-460-269C-6  from: 1  to: 922

      4  GTCCACGGCGACCGCAAGAAACCGCATTTGGCTAGCGATGCCATCGG  53
          :::::| | | | | :::::| | | | | | | | | | | | | | |
      875  LeuArgGlyThrArgAlaGluLeuGlyLeuGlyMetAlaAlaLeuGI  891
          | | | | | :::::| | | | | | | | | | | | | | |
      54  CCTTGGCTGATCTTTACCTGTGCGAAGCTATGAATAT  93
          | | | | | :::::| | | | | | | | | | | | | | |
      891  YArgGly.....HisSerLeuTyrAlaSerTyrGluTyr  902
;
```





```

|||||
1 1LevalHisGlyAspArglyLeuThrAlaIleGlyLeuAlaIleAlaIle 17
51 CGGCGCTGGTGGATCTTACCCGTGCTCCAGCCCTATGAATATTATGAAA 100
17 eGlyLeuGlyLeuTyrPheThrLeuGlyGlnAlaTyrGluTyrGluIle 34
101 TCGTCATACCGCAA 114
|||||
34 1LevalHisThrGlu 38

```

seq\_name: /SIDS2/gcgdata/geneseq/geneseqp/AA2001.DAT: AAB63840

seq\_documentation\_block:

ID AAB63840 standard; Protein: 93 AA.

AC AAB63840;

DT 26-MAR-2001 (first entry)

DE Human prostate cancer associated antigen protein sequence SEQ ID NO:1202.

KW Human; breast cancer; gastric cancer; prostate cancer; diagnosis;

KW cancer associated antigen; cytostatic; cancer vaccine.

OS Homo sapiens.

PN WO200073801-A2.

PD 07-DEC-2000.

PF 26-MAY-2000; 2000WO-US14749.

PR 28-MAY-1999; 99US-0136526.

PR 10-SEP-1999; 99US-0153454.

PA (LUDM-) LUDWIG INST CANCER RES.

PI Obata Y;

DR WPI: 2001-025274/03.

PT Nucleic acids encoding breast, gastric and prostate cancer associated

PT antigen precursors, useful for diagnosing and treating a condition

PT characterized by expression of an abnormal amount of a protein, e.g.

PS Example 1; Page 741; 799pp; English.

CC AAF22422 to AAF22626, AAF22627 to AAF22773 and AAF22774 to AAF23014

CC represent nucleotide sequences encoding human breast, gastric and

CC prostate cancer associated antigen precursors (CAAP) respectively.

CC AAB63232 to AAB63467, AAB63468 to AAB63721 and AAB63722 to AAB63970

CC represent human breast, gastric and prostate CAAP protein sequence

CC respectively. CAAPs have cytostatic activity and can be used in the

CC production of cancer vaccines. The human CAAP proteins, peptides, nucleic

CC acids or anti-CAAP antibodies are useful for diagnosing and treating a

CC condition characterized by expression of an abnormal amount of a protein,

CC e.g. cancer.

Sequence 93 AA;

alignment\_scores:

Quality: 62.00

Ratio: 2.583

Percent Similarity: 72.727

Percent Identity: 42.424

alignment\_block:

US-09-712-768-5 x AAB63840

Align seg 1/1 to: AAB63840 from: 1 to: 93

```

1 ATGCTCCAGCGGACCGCAGAAAACCCGATGTGCGCTAGCGATTCGCAT 50
5 LeuIleGluAsnAsnArgAsnGlnIleIleGlnAlaLeuLeuIleThrIle 21
51 CGGCGCTGGTGGATCTTACCCGTGCTCCAGCCCTATGAATATTATGAAA 99
21 eLeuLeuGlyLeuTyrPheThrLeuGlyGlnAlaSerGluTyr-PheGlu 37

```

seq\_name: /SIDS2/gcgdata/geneseq/geneseqp/AA2000.DAT: AAB29910

seq\_documentation\_block:

ID AAB29910 standard; Protein: 133 AA.

AC AAB29910;

DT 09-FEB-2001 (first entry)

DE Human secreted protein BLAST search protein SEQ ID NO: 168.

KW Cytostatic; immunosuppressive; nocotropic; neuroprotective; antiviral;

KW anti-allergic; hepatotropic; antidiabetic; anti-inflammatory; anticancer;

KW vanguard; anticonvulsant; antibacterial; antifungal; antiparasitic;

KW neurologic disease; infection; human; secreted protein.

OS Homo sapiens.

PN WO200061779-A1.

PD 19-OCT-2000.

PF 06-APR-2000; 2000WO-US09068.

PR 09-APR-1999; 99US-0128699.

PR 20-JAN-2000; 2000US-0177050.

PA (HUMA-) HUMAN GENOME SCI INC.

PI Rosen CA, Ruben SM, Komatsoulis G;

DR WPI: 2000-647424/62.

PT Isolated nucleic acid molecule encoding a human secreted protein is

PT used in preventing, treating or ameliorating a medical condition

PS Disclosure; Page 488-489; 495pp; English.

CC The invention relates to the isolation of genes AAC63410-C63458 encoding

CC the human secreted proteins AAB29802-B29850. This sequence represents a

CC line. The sequence is used as a query sequence for doing BLASTX searches

CC to determine homologous sequence to the protein. The genes and proteins

CC are useful for preventing, ameliorating or treating medical conditions,

CC e.g. by protein or gene therapy. The genes are isolated from a range of

CC human tissues disclosed in the specification. The genes are isolated from a range of

CC proteins, antibodies and (ant)agonists are useful in the diagnosis,

CC treatment and prevention of: (a) cancer, e.g. breast and ovarian cancer,

CC and other cancers of the adrenal gland, bone, bone marrow, breast,

CC gastrointestinal tract, liver, lung, or urogenital; (b) immune disorders

CC e.g. Addison's disease, allergies, autoimmune haemolytic anaemia,

CC autoimmune thyroiditis, diabetes mellitus, Crohn's disease, multiple

CC sclerosis, rheumatoid arthritis and ulcerative colitis; (c) multiple

CC cardiovascular disorders such as myocardial ischaemias; (d) wound

CC healing; (e) neurological diseases e.g. cerebral anoxia and epilepsy; and

CC (f) infectious diseases such as viral, bacterial, fungal and parasitic

Sequence 133 AA;

alignment\_scores:

Quality: 62.00

Ratio: 2.583

Length: 33

Gaps: 0

Percent Similarity: 72.727 Percent Identity: 42.424

alignment\_block:

US-09-712-768-5 x AAB75604 ..

Align seg 1/1 to: AAB75604 from: 1 to: 133

```

1 ATGCTCCAGCGCGCAAGAAACCGCGATTGGCTTATGATGCAAT 50
   : : : : : : : : : : : : : : : : : : : : : : : :
23 LeuLeuGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGln 39

```

```

51 CGGCTTGGCTGATCTTACCTGTCGCAAGCTTATGATGCAAT 99
   | | | | | | | | | | | | | | | | | | | | | | |
39 eLeuLeuGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGln 55

```

seq.name: /SID52/gcgdata/geneseq/geneseq/AA2001.DAT: AAB75604

seq\_documentation\_block:

ID AAB75604 standard; Protein: 203 AA.

AC AAB75604;

DT 06-Apr-2001 (first entry)

Human secreted protein sequence encoded by gene 46 SEQ ID NO:158.

Human: secreted protein; immunosuppressive; antiarthritic; antirheumatic; antiproliferative; cytostatic; cardiac; vasotropic; cerebroprotective; neuroprotective; antibacterial; virucide; fungicide; opthalmological; vulnary; autoimmune disease; cardiovascular disorder; hyperproliferative disorders; cerebrovascular disorder; wound healing; nervous system disorder; ocular disorder; skin ageing; chemotaxis; food additive.

OS Homo sapiens.

PN WO200077026-A1.

PD 21-DEC-2000.

PF 01-JUN-2000; 2000WO-US14973.

PR 11-JUN-1999; 99US-0138630.

PA (HUMA-) HUMAN GENOME SCI INC.

PI (ROSE/) ROSEN C A.

PI Rosen CA, Ruben SM, Komatsoulis GA;

PI WPI: 2001-071258/08.

PI N-PSDB: AAF64221.

PI Nucleic acid molecules encoding human secreted proteins, used in preventing, treating or ameliorating a disorder, e.g. Alzheimer's and Parkinson's diseases and cancers -

PI Disclosure: Page 94; 542pp; English.

Human secreted proteins AAB75506 - AAB75554 are encoded by polynucleotide sequences AAB75506 - AAB75554. The specification includes amino acid sequences AAB75506 - AAB75554 which represent fragments of the human secreted proteins, and protein sequences with which they share homology. The proteins and polynucleotides, their agonists and antagonists have activities dependent on the tissues and cells in which they are expressed, examples of these activities include: immunosuppressive; antiarthritic; antirheumatic; antiproliferative; cytostatic; cardiac; vasotropic; cerebroprotective; neuroprotective; antibacterial; virucide; fungicide; opthalmological; and vulnary. The proteins, polynucleotides, agonists and antagonists can be used to treat or detect diseases e.g. rheumatoid arthritis, hyperproliferative disorders e.g. neoplasms of the breast or liver, cardiovascular disorders e.g. cardiac arrest, cerebrovascular disorders e.g. cerebral ischaemia,

anglogenesis, nervous system disorders e.g. Alzheimer's disease, infections caused by bacteria, viruses and fungi and ocular disorders e.g. corneal infection. The polypeptides can also be used to aid wound healing and epithelial cell proliferation, to prevent skin ageing due to sunburn, to maintain organs before transplantation, for supporting cell culture of primary tissues, to regenerate tissues and in chemotaxis. The polypeptides can also be used as a food additive or preservative to increase or decrease storage capabilities. Included in the invention are polynucleotide sequences AAB75506 - AAB75554 and peptide AAB75505 which are used in the isolation, identification and characterisation of the proteins of the invention.

XX Sequence 203 AA;

alignment\_scores: Length: 33

Quality: 62.00

Ratio: 2.583

Percent Similarity: 72.727 Percent Identity: 42.424

alignment\_block:

US-09-712-768-5 x AAB75604 ..

Align seg 1/1 to: AAB75604 from: 1 to: 203

```

1 ATGCTCCAGCGCGCAAGAAACCGCGATTGGCTTATGATGCAAT 50
   : : : : : : : : : : : : : : : : : : : : : : :
93 LeuLeuGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGln 109

```

```

51 CGGCTTGGCTGATCTTACCTGTCGCAAGCTTATGATGCAAT 99
   | | | | | | | | | | | | | | | | | | | | | | |
109 eLeuLeuGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGln 125

```

seq.name: /SID52/gcgdata/geneseq/geneseq/AA1999.DAT: AAB75610

seq\_documentation\_block:

ID AAB75610 standard; Protein: 455 AA.

AC AAB75610;

DT 13-SEP-1999 (first entry)

Human secreted protein sequence.

Human: secreted protein; pneumonia; bronchitis; heart disease; sarcoidosis; respiratory disease; pneumonia; bronchitis; heart disease; sarcoidosis; sinusitis; purulent otitis media; erythema nodosum; pharyngitis; vaccine; neutralising epitope.

OS Chlamydia pneumoniae.

PN WO9927105-A2.

PD 03-JUN-1999.

PF 20-NOV-1998; 98WO-1B01890.

PR 04-NOV-1998; 98US-0107078.

PR 21-NOV-1997; 97FR-0014673.

PA (GEST ) GENSET.

PI Griffais R;

PI WPI: 1999-357842/30.

PI Genome sequence of Chlamydia pneumoniae

PI Page 1338-1339; Disclosure: 1912pp; English.

AA734584-Y35879 represent the proteins encoded by all the open reading frames in the complete genome (see AA734584) of Chlamydia pneumoniae. C. pneumoniae causes respiratory disease such as pneumonia and bronchitis and is thought to be a contributing factor in heart

CC disease, sarcoidosis, sinusitis, purulent otitis media, erythema  
 CC nodosum or pharyngitis. The polypeptides encoded by the open reading  
 CC frames of the C. pneumoniae genome (see AAY34584-135879) can be used in  
 CC immunogenic compositions as vaccines. Vectors containing C. pneumoniae  
 CC nucleotide sequences can also be used as immunogenic compositions,  
 CC especially where the vector directs the expression of a neutralising  
 CC epitope of C. pneumoniae.  
 XX  
 SO Sequence 455 AA.

## alignment\_scores:

Quality: 62.00 Length: 18  
 Ratio: 4.133 Gaps: 0  
 Percent Similarity: 83.333 Percent Identity: 50.000

## alignment\_block:

US-09-712-768-5 x AAY35610 ..

Align seg 1/1 to: AAY35610 from: 1 to: 455

61 TGGATCTTACCCCTGCGCAGCCATGATATATGAAATCGTCATAC 110  
 |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:  
 35 TTPLeuTYrThrLeuCYGlnSerTyrGlnHisLysLysLeuValPhePr 51  
 111 CGAA 114  
 |||  
 51 GGLu 52

seq\_name: /SID2/gcgdata/geneseq/geneseq/AA2000.DAT: AAB29909  
 seq\_documentation\_block:  
 ID AAB29909 standard; Protein: 133 AA.

AC AAB29909;

DT 09-FEB-2001 (first entry)

DE Human secreted protein BLAST search protein SEQ ID NO: 167.

XX Cytostatic; immunosuppressive; nootropic; neuroprotective; antiviral;  
 KW antiallergic; hepatotropic; antidiabetic; antiinflammatory; antitumor;  
 KW vulnerability; anticonvulsant; antibacterial; antifungal; antiparasitic;  
 KW cardiac; gene therapy; cancer; immune disorder; cardiovascular disorder;  
 KW neurological disease; infection; human; secreted protein.  
 OS Homo sapiens.

PN WO200061779-A1.

PD 19-OCT-2000.

PE 06-APR-2000; 2000WO-US09068.

PR 09-APR-1999; 99US-0128699.

PR 20-JAN-2000; 2000US-0177030.

PA (HUMA-) HUMAN GENOME SCI INC.

PI Rosen CA, Ruben SM, Komatsoulis G;

DR WPI: 2000-647424/62.

XX Isolated nucleic acid molecule encoding a human secreted protein is  
 PT used in preventing, treating or ameliorating a medical condition  
 XX  
 PS Disclosure: Page 487; 495pp; English.

CC The invention relates to the isolation of genes AAC63410-C63458 encoding  
 CC the human secreted proteins AAB29802-B29850. This sequence represents a  
 CC peptide fragment homologous to the protein encoded by the gene given  
 CC in the descriptor line. The sequence is a search result from a BLASTX  
 CC homology search. The genes and proteins are useful for preventing,

CC ameliorating or treating medical conditions, e.g. by protein or gene  
 CC therapy. The genes are isolated from a range of human tissues disclosed  
 CC in the specification. The nucleic acids, proteins, antibodies and  
 CC (ant)agonists are useful in the diagnosis, treatment and prevention of:  
 CC (a) cancer, e.g. breast and ovarian cancer, and other cancers of the  
 CC adrenal gland, bone, bone marrow, breast, gastrointestinal tract, liver,  
 CC lung, or urogenital; (b) immune disorders e.g. Addison's disease,  
 CC allergies, autoimmune haemolytic anaemia, autoimmune thyroiditis,  
 CC diabetes mellitus, Crohn's disease, multiple sclerosis, rheumatoid  
 CC arthritis and ulcerative colitis; (c) cardiovascular disorders such  
 CC as myocardial ischaemia; (d) wound healing; (e) neurological diseases  
 CC e.g. cerebral anoxia and epilepsy; and (f) infectious diseases such  
 CC as viral, bacterial, fungal and parasitic infections.  
 XX  
 SO Sequence 133 AA.

## alignment\_scores:

Quality: 60.00 Length: 33  
 Ratio: 2.500 Gaps: 0  
 Percent Similarity: 72.727 Percent Identity: 42.424

## alignment\_block:

US-09-712-768-5 x AAB29909 ..

Align seg 1/1 to: AAB29909 from: 1 to: 133

1 ATGTCACAGCGCGCAGGAGAAACCGCGATGCGCTACGATTCGAT 50  
 :|||:|||||:|||||:|||||:|||||:|||||:|||||:  
 23 LeuMetGluAsnArgAsnGlnMetCLeuAlaLeuLeuLeuThrLl 39  
 51 CGGCTTGCGCTGATCTTACCTGTGCGCAGCCATGAAATATATGAA 99  
 | ||||| ||||| ||||| ||||| ||||| ||||| |||||  
 39 eLeuLeuGlyLeuTyrPheThrLeuLeuGlnAlaSerGluTyrPheGlu 55

seq\_name: /SID2/gcgdata/geneseq/geneseq/AA2000.DAT: AAB36196

seq\_documentation\_block:  
 ID AAB36196 standard; Protein: 172 AA.

AC AAB36196;

DT 02-MAR-2001 (first entry)

DE Cowdria ruminantium 4hworf1 polypeptide.

XX Cowdria ruminantium; MAP1; major antigenic protein 1; antirickettsial;  
 KW vaccine; gene therapy; Rickettsia; Cowdria; Anaplasma; map2; ihworf3;  
 KW 4hworf1; 18hworf1; 3gdorf3.

OS Cowdria ruminantium.

PN WO200065063-A2.

PD 02-NOV-2000.

PE 21-APR-2000; 2000WO-US10886.

PR 22-APR-1999; 99US-0130725.

PA (UYFL) UNIV FLORIDA.

PI Barbet AF, Bowie MV, Gantla RR, Burridge MJ, Mahan SM, McGuire TC;  
 PI Rurangirwa FR, Moreland AL, Simbi BH, Whitmore MW, Alleman AK;

DR WPI: 2000-679675/66.  
 DR N-PSDB; AAC68713.

XX New polynucleotides useful as DNA vaccines for conferring immunity to  
 PT rickettsial infection e.g. heartwater caused by Cowdria ruminantium,  
 PT comprises major antigenic protein 1 or 2 gene of rickettsial pathogens







PR 29-APR-1998; 98US-0083545.  
 PR 29-APR-1998; 98US-0083554.  
 PR 29-APR-1998; 98US-0083558.  
 PR 29-APR-1998; 98US-0083559.  
 PR 30-APR-1998; 98US-0083742.  
 PR 05-MAY-1998; 98US-0083466.  
 PR 06-MAY-1998; 98US-0084414.  
 PR 06-MAY-1998; 98US-0084411.  
 PR 07-MAY-1998; 98US-0084598.  
 PR 07-MAY-1998; 98US-0084600.  
 PR 07-MAY-1998; 98US-0084627.  
 PR 07-MAY-1998; 98US-0084637.  
 PR 07-MAY-1998; 98US-0084639.  
 PR 07-MAY-1998; 98US-0084640.  
 PR 07-MAY-1998; 98US-0084643.  
 PR 13-MAY-1998; 98US-0085323.  
 PR 13-MAY-1998; 98US-0085338.  
 PR 13-MAY-1998; 98US-0085339.  
 PR 15-MAY-1998; 98US-0085573.  
 PR 15-MAY-1998; 98US-0085579.  
 PR 15-MAY-1998; 98US-0085580.  
 PR 15-MAY-1998; 98US-0085582.  
 PR 15-MAY-1998; 98US-0085689.  
 PR 15-MAY-1998; 98US-0085697.  
 PR 15-MAY-1998; 98US-0085700.  
 PR 15-MAY-1998; 98US-0085704.  
 PR 18-MAY-1998; 98US-0086023.  
 PR 22-MAY-1998; 98US-0086392.  
 PR 22-MAY-1998; 98US-0086414.  
 PR 22-MAY-1998; 98US-0086430.  
 PR 28-MAY-1998; 98US-0086486.  
 PR 28-MAY-1998; 98US-0087096.  
 PR 28-MAY-1998; 98US-0087106.  
 PR 30-JUL-1998; 98US-0087208.  
 PR 11-SEP-1998; 98US-0094651.  
 PR 11-SEP-1998; 98US-0100038.

XX (GETH ) GENENTECH INC.

XX Wood WI, Goddard A, Gurney A, Yuan J, Baker KP, Chen J;

XX WPI; 1999-551358/46.

DR N-PSDB; AA234275.

XX New secreted and transmembrane polypeptides and their polynucleotides,  
 PT useful for treating blood coagulation disorders, cancers and cellular  
 PT adhesion disorders -

XX Claim 12; Fig 190; 530pp; English.

XX The present invention describes secreted and transmembrane polypeptides  
 CC and their polynucleotides. The nucleotide sequences are useful as  
 CC sources of probes, primers, for chromosome mapping, and for generation  
 CC of antisense sequences. They can also be used to create transgenic  
 CC animals. The proteins can be used to treat a variety of diseases and  
 CC disorders, depending on their function. Diseases that may be treated  
 CC include blood coagulation disorders, cancers and cellular adhesion  
 CC disorders. They may also be used to raise antibodies. AA23891 to  
 CC AA23433, and AA41685 to AA41774 represent polynucleotide and  
 CC polypeptide sequence given in the exemplification of the present  
 CC invention.

XX Sequence 747 AA;

SO alignment\_scores:

Quality: 58.00 Length: 48  
 Ratio: 2.762 Gaps: 1

Percent Similarity: 43.750 Percent Identity: 27.083

alignment\_block:

US-09-712-768-5 x AA41758

Align seg 1/1 to: AA41758 from: 1 to: 747

7 CACGGCGACCGCAGAAA..... 24  
 217 HSGLYASPARGSErLYSGluserLeuValserPheAlaMetGlnHsVa 233  
 25 .....ACCGGATG 34  
 233 lARSerThrValThrGlueTrpThrGlyAsnPhaValasnSerIleG 250  
 35 GCCTAGCATGTCACATCGCGCTTGCGTGCATCTTACCCGTGC 78  
 250 lNThrAlaPheAlaIacGlyIleGlyTrpLeuIleThrPheCys 264  
 seq\_name: /SIDS2/gcgdata/geneseq/geneseq/AA2000.DAT:AA44314

seq\_documentation\_block:

ID AAB44314 standard; Protein: 747 AA.

AC AAB44314;

DT 08-FEB-2001 (first entry)

DE Human PRO1012 (UNQ495) protein sequence SEQ ID NO:459.

XX Human: secreted protein; transmembrane protein; PRO; EST; cytosolic;  
 KW expressed sequence tag; detection; cancer.

XX Homo sapiens.

PN W0200053756-A2.

XX 14-SEP-2000.

PD 18-FEB-2000; 2000WO-US04311.

PF 08-MAR-1999; 99WO-US05028.

PR 12-MAR-1999; 99US-0123957.

PR 29-MAR-1999; 99US-0126773.

PR 21-APR-1999; 99US-0130232.

PR 28-APR-1999; 99US-0131445.

PR 14-MAY-1999; 99US-0134287.

PR 23-JUN-1999; 99US-0141037.

PR 26-JUL-1999; 99US-0145698.

PR 29-OCT-1999; 99US-0162506.

PR 30-NOV-1999; 99WO-US28313.

PR 02-DEC-1999; 99WO-US28351.

PR 16-DEC-1999; 99WO-US28565.

PR 30-DEC-1999; 99WO-US31243.

PR 05-JAN-2000; 2000WO-US00219.

PR 06-JAN-2000; 2000WO-US00277.

PR 06-JAN-2000; 2000WO-US00376.

XX (GETH ) GENENTECH INC.

XX Ashkenazi AJ, Baker KP, Botstein D, Desnoyers L, Eaton DL;

XX Ferrara N, Filvaroff E, Fong S, Gao W, Gerber H, Gerritsen ME;  
 XX Goddard A, Godowski PJ, Grimaldi CJ, Gurney AL, Hillan KJ;  
 XX Kljavin IJ, Kuo SS, Napier MA, Pan J, Paoni NF, Roy MA;  
 XX Shelton DL, Stewart RA, Tumas D, Williams PM, Wood WI;  
 XX WPI; 2000-611443/58.  
 XX N-PSDB; AAC78573.  
 XX Novel PRO polypeptides and polynucleotides used in detection methods,  
 PT to target bioactive molecules to specific cells, and to modulate  
 PT cellular activities -  
 XX Claim 12; Fig 190; 636pp; English.  
 XX AAC78458 to AAC78599 represent polynucleotide and EST (expressed

CC sequence tag) sequences which encode secreted or transmembrane PRO  
 CC polypeptides. The PRO polynucleotides and polypeptides have cytostatic  
 CC activity. The polynucleotides and polypeptides can be used for detecting  
 CC the presence of PRO polypeptides in samples, for linking bioactive  
 CC molecules to cells and for modulating biological activities of cells,  
 CC using the polypeptides for specific targeting. The polypeptide targeting  
 CC can be used to kill the target cells, e.g. for the treatment of cancers.  
 CC The polypeptide pairs provide specific targeting of bioactive molecules  
 CC to cells. AAC78600 to AAC78987 represent PCR primers and probes used in  
 CC the isolation of the PRO polynucleotide sequences.

Sequence 747 AA;

alignment\_scores:

Quality: 58.00 Length: 48  
 Ratio: 2.762 Gaps: 1  
 Percent Similarity: 43.750 Percent Identity: 27.083

alignment\_block:

US-09-712-768-5 x AAB44314 ..

Align seg 1/1 to: AAB44314 from: 1 to: 747

7 CACGGCAGCCGACAGAAA..... 24  
 |||||  
 217 HisGlyAspArgSerLysGluSerLeuValSerPheAlaMetGlnHisVa 233  
 25 .....ACCGGATTG 34  
 223 LArgSerThrValThrGluLeuTrpHnGlyAsnPheValAsnSerIleG 250  
 35 GCCAGGATTCGATCGCCTTGCTGGATCTTACCCCTGTC 78  
 |||||  
 250 LnrHnAlaPheAlaIleGlyIleGlyTrpLeuIleTrpHnPhCyS 264

seq\_name: /SIDS2/gcdata/geneseq/geneseq/AA2001.DAT:AAU12384

seq\_documentation\_block:

ID AAU12384 standard; Protein; 747 AA.

XX AAU12384;  
 AC  
 DT 24-OCT-2001 (first entry)  
 XX  
 DE Human PRO1012 polypeptide sequence.  
 XX  
 KW Human secretory and transmembrane; PRO; mammalian; cancer; lung;  
 KW breast; prostate; cervical; tumor necrosis factor-alpha; TNF-alpha;  
 KW cartilage; ear; proliferation; glucose; free fatty acid; skeletal muscle;  
 KW adipocyte; A-peptide; factor VIIA; gene therapy.  
 XX  
 OS Homo sapiens.  
 XX  
 PN WO200140466-A2.  
 PD  
 XX  
 PD 07-JUN-2001.  
 XX  
 PF 01-DEC-2000; 2000WO-US32678.  
 XX  
 XX  
 PR 01-DEC-1999; 99WO-US28301.  
 PR 01-DEC-1999; 99WO-US28634.  
 PR 02-DEC-1999; 99WO-US28551.  
 PR 02-DEC-1999; 99WO-US28564.  
 PR 02-DEC-1999; 99WO-US28565.  
 PR 09-DEC-1999; 99US-0170262.  
 PR 16-DEC-1999; 99WO-US30095.  
 PR 20-DEC-1999; 99WO-US30911.  
 PR 20-DEC-1999; 99WO-US30993.  
 PR 30-DEC-1999; 99WO-US31243.  
 PR 06-JAN-2000; 2000WO-US00277.  
 PR 06-JAN-2000; 2000WO-US00376.  
 PR 11-FEB-2000; 2000WO-US03565.

PR 18-FEB-2000; 2000WO-US04341.  
 PR 18-FEB-2000; 2000WO-US04342.  
 PR 22-FEB-2000; 2000WO-US04414.  
 PR 24-FEB-2000; 2000WO-US04914.  
 PR 24-FEB-2000; 2000WO-US05004.  
 PR 01-MAR-2000; 2000WO-US05601.  
 PR 20-MAR-2000; 2000WO-US07377.  
 PR 21-MAR-2000; 2000WO-US07532.  
 PR 30-MAR-2000; 2000WO-US08439.  
 PR 17-MAY-2000; 2000WO-US13705.  
 PR 22-MAY-2000; 2000WO-US14042.  
 PR 30-MAY-2000; 2000WO-US14941.  
 PR 02-JUN-2000; 2000WO-US15264.  
 PR 10-NOV-2000; 2000WO-US30873.

(GETH ) GENENTECH INC.

PI Baker KP, Beresini M, Deforge L, Desnoyers L, Filvaroff E, Gao W;  
 PI Gerritsen ME, Goddard A, Godowski PJ, Gurney AL, Sherwood S;  
 PI Smith V, Stewart RA, Tumas D, Watanabe CK, Wood WI, Zhang Z;  
 XX  
 DR WPI: 2001-408281/43.  
 DR N-PSDB; AAS21456.

XX  
 PT Isolated, secretory and transmembrane PRO polypeptide used to detect  
 PT other PRO polypeptides, link bioactive molecules to cells expressing  
 PT PRO polypeptides, and detect the presence of mammalian tumours e.g.  
 PT lung, breast, prostate, cervical  
 XX  
 PS Claim 12; Fig 426; 813pp; English.

XX  
 CC AAU12172-AAU12446 represent novel human secretory and transmembrane  
 CC PRO polypeptides. The PRO polypeptides are useful to detect other  
 CC PRO polypeptides, to link bioactive molecules to cells expressing  
 CC PRO polypeptides, to modulate biological activities of cells expressing  
 CC PRO polypeptides, and to detect the presence of mammalian lung, colon,  
 CC breast, prostate, rectal, cervical or liver tumours by comparing PRO  
 CC polypeptide expression in a cell sample to that in a control sample.  
 CC Some of the 275 sequences are also useful to stimulate the release of  
 CC tumor necrosis factor-alpha (TNF-alpha) from human blood, the  
 CC proliferation or differentiation of chondrocytes, the proliferation or  
 CC gene expression in pericyte cells, the release of proteoglycans from  
 CC cartilage, the proliferation of inner ear utricular supporting cells or  
 CC of T-lymphocytes, the release of a cytokine from peripheral blood  
 CC monocytes (PBMCs), or the proliferation of endothelial cells. Some of  
 CC the PRO polypeptides may modulate glucose or free fatty acid uptake by  
 CC skeletal muscle cells or by adipocytes, or inhibit binding of A-peptide  
 CC to factor VIIA. The PRO polypeptides can be used in assays to identify  
 CC molecules involved in binding interactions. The polynucleotides encoding  
 CC PRO polypeptides can be used to generate probes, antisense RNA/DNA,  
 CC transgenic or knock out animals and can be used in gene therapy.

Sequence 747 AA;

alignment\_scores:

Quality: 58.00 Length: 48  
 Ratio: 2.762 Gaps: 1  
 Percent Similarity: 43.750 Percent Identity: 27.083

alignment\_block:

US-09-712-768-5 x AAU12384 ..

Align seg 1/1 to: AAU12384 from: 1 to: 747

7 CACGGCAGCCGACAGAAA..... 24  
 |||||  
 217 HisGlyAspArgSerLysGluSerLeuValSerPheAlaMetGlnHisVa 233  
 25 .....ACCGGATTG 34  
 223 LArgSerThrValThrGluLeuTrpHnGlyAsnPheValAsnSerIleG 250

35 GCCTAGCGATTGCCATCGCCTTGCGTACCTTACCTGTGC 78  
 |||::||| |||::|||::|||::|||::|||  
 250 lnrhAlaPheAlaAlaGlylIeglyTrpLeuIleThrPheCys 264

seq\_name: /SID52/gcgdata/geneseq/geneseqp/AA2001.DAT: AAB92743

seq\_documentation\_block:

ID AAB92743 standard; Protein; 768 AA.

AC AAB92743;

DT 26-JUN-2001 (first entry)

DE Human protein sequence SEQ ID NO:11190.

DE Human; primer; detection; diagnosis; antisense therapy; gene therapy.

OS Homo sapiens.

PN EPI074617-A2.

PD 07-FEB-2001.

PF 28-JUL-2000; 2000EP-0116126.

PR 29-JUL-1999; 99JP-0248036.

PR 27-AUG-1999; 99JP-0300253.

PR 11-JAN-2000; 2000JP-0118776.

PR 02-MAY-2000; 2000JP-0183767.

PR 09-JUN-2000; 2000JP-0241899.

PA (HELI-) HELIX RES INST.

PI Ota T, Isogai T, Nishikawa T, Hayashi K, Saito K, Yamamoto J;

PI Ishii S, Sugiyama T, Wakamatsu A, Nagai K, Otsuki T;

DR WPI; 2001-318749/34.

PT Primer sets for synthesizing polynucleotides, particularly the 5602

PT full-length cDNAs defined in the specification, and for the detection

PT and/or diagnosis of the abnormality of the proteins encoded by the

PT full-length cDNAs -

PS Claim 8; SEQ ID 11190; 2537pp + CD ROM; English.

XX The present invention describes primer sets for synthesizing 5602

CC full-length cDNAs defined in the specification. Where a primer set

CC comprises: (a) an oligo-dT primer and an oligonucleotide complementary

CC to the complementary strand of a polynucleotide which comprises one of

CC the 5602 nucleotide sequences defined in the specification, where the

CC oligonucleotide comprises at least 15 nucleotides; or (b) a combination

CC of an oligonucleotide comprising a sequence complementary to the

CC sequence and an oligonucleotide comprising a sequence complementary to a

CC polynucleotide which comprises a 5'-end

CC sequence and an oligonucleotide comprising a 3'-end sequence, where the

CC oligonucleotide comprises at least 15 nucleotides and the combination of

CC the 5'-end sequence/3'-end sequence is selected from those defined in

Ratio: 2.762 Gaps: 1  
 Percent Similarity: 43.750 Percent Identity: 27.083

alignment\_block:

US-09-712-768-5 x AAB92743 ..

Align seq 1/1 to: AAB92743 from: 1 to: 768

7 CACGCGCAGCCGAGAGAA..... 24

192 HisGlySPARGSerLysGluSerLeuValSerPheAlaMetGlnHisVa 208

25 .....ACCGCATG 34

208 lArSerThrValThrGluLeuTrpThrGlyAsnPheValAsnSerIleG 225

35 GCCTAGCGATTGCCATCGCCTTGCGTACCTTACCTGTGC 78

225 lnrhAlaPheAlaAlaGlylIeglyTrpLeuIleThrPheCys 239

seq\_name: /SID52/gcgdata/geneseq/geneseqp/AA2001.DAT: AAB93204

seq\_documentation\_block:

ID AAB93204 standard; Protein; 793 AA.

AC AAB93204;

DT 26-JUN-2001 (first entry)

DE Human protein sequence SEQ ID NO:12172.

DE Human; primer; detection; diagnosis; antisense therapy; gene therapy.

OS Homo sapiens.

PN EPI074617-A2.

PD 07-FEB-2001.

PF 28-JUL-2000; 2000EP-0116126.

PR 29-JUL-1999; 99JP-0248036.

PR 27-AUG-1999; 99JP-0300253.

PR 11-JAN-2000; 2000JP-0118776.

PR 02-MAY-2000; 2000JP-0183767.

PR 09-JUN-2000; 2000JP-0241899.

PA (HELI-) HELIX RES INST.

PI Ota T, Isogai T, Nishikawa T, Hayashi K, Saito K, Yamamoto J;

PI Ishii S, Sugiyama T, Wakamatsu A, Nagai K, Otsuki T;

DR WPI; 2001-318749/34.

PT Primer sets for synthesizing polynucleotides, particularly the 5602

PT full-length cDNAs defined in the specification, and for the detection

PT and/or diagnosis of the abnormality of the proteins encoded by the

PT full-length cDNAs -

PS Claim 8; SEQ ID 12172; 2537pp + CD ROM; English.

XX The present invention describes primer sets for synthesizing 5602

CC full-length cDNAs defined in the specification. Where a primer set

CC comprises: (a) an oligo-dT primer and an oligonucleotide complementary

CC to the complementary strand of a polynucleotide which comprises one of

CC the specification. The primer sets can be used in antisense therapy and  
CC in gene therapy. The primers are useful for synthesizing polynucleotides,  
CC particularly full-length cDNAs. The primers are also useful for the  
CC detection and/or diagnosis of the abnormality of the proteins encoded by  
CC the full-length cDNAs. The primers allow obtaining of the full-length  
CC cDNAs easily without any specialised methods. AAH03166 to AAH13628 and  
CC AAH13633 to AAH18742 represent human cDNA sequences; AAB92446 to  
CC AAB9893 represent human amino acid sequences; and AAH13629 to AAH13632  
CC represent oligonucleotides, all of which are used in the exemplification  
CC of the present invention.

XX Sequence 793 AA;

alignment\_scores:  
Quality: 58.00 Length: 48  
Ratio: 2.762 Gaps: 1  
Percent Similarity: 43.750 Percent Identity: 27.083

alignment\_block:

US-09-712-768-5 x AAB93204 ..

Align seg 1/1 to: AAB93204 from: 1 to: 793

7 CACGGCGACCGCAGAGAA..... 24  
217 HisGlyAspArgSerLysGluSerLeuValSerPheAlaMetGlnHisVa 233  
25 .....ACCGGATG 34  
233 LarSerThrValThrGluLeuThrPthrGlyAsnPheValAsnSerIleG 250  
35 GCCTAGCGATTGCGCCTGCGTGCATCTTTACCGCTGTC 78  
250 InThrAlaPheAlaIaGlyIleGlyTrpLeuIleThrPheCys 264

seq\_name: /SID2/gcdata/geneseq/geneseq/AAB9309

seq\_documentation\_block:

ID AAB95309 standard; Protein: 793 AA.

AC AAB95309;  
XX  
XX 26-JUN-2001 (first entry)  
XX  
XX Human protein sequence SEQ ID NO:17552.  
DE  
XX Human; primer; detection; diagnosis; antisense therapy; gene therapy.  
XX  
XX Homo sapiens.  
XX  
XX PN EPI074617-A2.  
XX  
XX 07-FEB-2001.  
XX  
XX 28-JUL-2000; 2000EP-0116126.  
XX  
XX 29-JUL-1999; 99JP-0248036.  
XX  
XX 27-AUG-1999; 99JP-0300253.  
XX  
XX 11-JAN-2000; 2000JP-0118776.  
XX  
XX 02-MAY-2000; 2000JP-0183767.  
XX  
XX 09-JUN-2000; 2000JP-0241899.  
XX  
XX  
XX (HELI-) HELIX RES INST.

XX Ota T, Isogai T, Nishikawa T, Hayashi K, Saito K, Yamamoto J;  
XX Ishii S, Sugiyama T, Wakamatsu A, Nagai K, Otsuki T;  
XX  
XX WPI; 2001-318749/34.

XX Primer sets for synthesizing polynucleotides, particularly the 5602  
XX full-length cDNAs defined in the specification, and for the detection  
XX and/or diagnosis of the abnormality of the proteins encoded by the

PT full-length cDNAs -  
XX  
XX Claim 8; SEQ ID 17552; 2537bp + CD ROM; English.

CC The present invention describes primer sets for synthesizing 5602  
CC full-length cDNAs defined in the specification. Where a primer set  
CC comprises: (a) an oligo-dr primer and an oligonucleotide complementary  
CC to the complementary strand of a polynucleotide which comprises one of  
CC the 5602 nucleotide sequences defined in the specification, where the  
CC oligonucleotide comprises at least 15 nucleotides; or (b) a combination  
CC of an oligonucleotide comprising a sequence complementary to the  
CC complementary strand of a polynucleotide which comprises a 5'-end  
CC sequence and an oligonucleotide comprising a sequence complementary to a  
CC polynucleotide which comprises a 3'-end sequence, where the  
CC oligonucleotide comprises at least 15 nucleotides and the combination of  
CC the 5'-end sequence/3'-end sequence is selected from those defined in  
CC the specification. The primer sets can be used in antisense therapy and  
CC in gene therapy. The primers are useful for synthesizing polynucleotides,  
CC particularly full-length cDNAs. The primers are also useful for the  
CC detection and/or diagnosis of the abnormality of the proteins encoded by  
CC the full-length cDNAs. The primers allow obtaining of the full-length  
CC cDNAs easily without any specialised methods. AAH03166 to AAH13628 and  
CC AAH13633 to AAH18742 represent human cDNA sequences; AAB92446 to  
CC AAB9893 represent human amino acid sequences; and AAH13629 to AAH13632  
CC represent oligonucleotides, all of which are used in the exemplification  
CC of the present invention.

XX Sequence 793 AA;

alignment\_scores:  
Quality: 58.00 Length: 48  
Ratio: 2.762 Gaps: 1  
Percent Similarity: 43.750 Percent Identity: 27.083

alignment\_block:

US-09-712-768-5 x AAB95309 ..

Align seg 1/1 to: AAB95309 from: 1 to: 793

7 CACGGCGACCGCAGAGAA..... 24  
217 HisGlyAspArgSerLysGluSerLeuValSerPheAlaMetGlnHisVa 233  
25 .....ACCGGATG 34  
233 LarSerThrValThrGluLeuThrPthrGlyAsnPheValAsnSerIleG 250  
35 GCCTAGCGATTGCGCCTGCGTGCATCTTTACCGCTGTC 78  
250 InThrAlaPheAlaIaGlyIleGlyTrpLeuIleThrPheCys 264